AD_____

GRANT NUMBER DAMD17-94-J-4124

TITLE: Molecular Recognition of Endocytic Codes in Receptor Tyrosine Kinases

PRINCIPAL INVESTIGATOR: Rui-Yun Wu
Dr. Gordon Gill

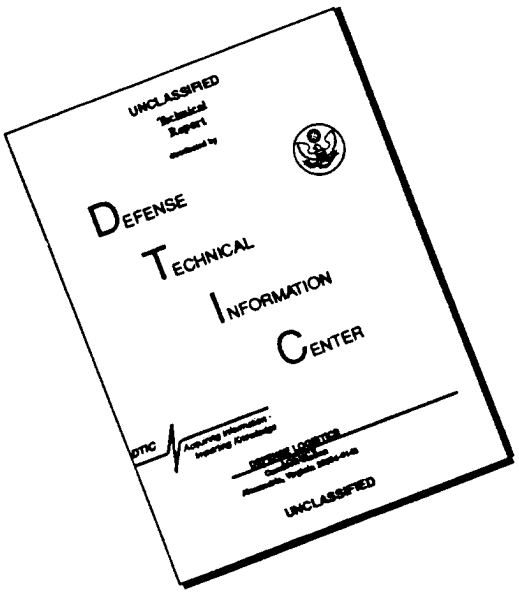
CONTRACTING ORGANIZATION: The University of California,
San Diego
La Jolla, California 92093

REPORT DATE: July 1996

TYPE OF REPORT: Final Report

PREPARED FOR: Commander
U.S. Army Medical Research and Materiel Command
Fort Detrick, Frederick, Maryland 21702-5012

DISCLAIMER NOTICE



THIS DOCUMENT IS BEST QUALITY AVAILABLE. THE COPY FURNISHED TO DTIC CONTAINED A SIGNIFICANT NUMBER OF PAGES WHICH DO NOT REPRODUCE LEGIBLY.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	July 1996	Final Report 1 Jul 94 - 30 Jun 96

4. TITLE AND SUBTITLE
Molecular Recognition of Endocytic Codes in Receptor Tyrosine Kinases

5. FUNDING NUMBERS
DAMD17-94-J-4124

6. AUTHOR(S)
Rui-Yun Wu
Dr. Gordon Gill

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
The University of California, San Diego
La Jolla, California 92093

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Commander
U.S. Army Medical Research and Materiel Command
Fort Detrick, MD 21702-5012

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

Enigma, isolated based on interaction with the human insulin receptor (InsR) using a yeast two-hybrid system, contains three LIM domains within its carboxyl terminus. The carboxyl LIM domain (LIM3) specifically recognizes the endocytic codes of InsR. Interaction of two random peptide libraries indicates the specific binding of Gly-Pro-Hyr-Gly-Pro-Hyr-Tyr/Phe corresponding to the major endocytic code of InsR. The ability of LIM3 of Enigma to recognize the endocytic codes of InsR fulfills the first property of the endocytic mechanism. In contrast to LIM3 of Enigma binding to InsR, LIM2 of Enigma associates specifically with the tyrosine kinase receptor Ret. Mutational analysis indicated that Tyr^{586} at the carboxyl terminus of Ret is essential for the Ret and Enigma interaction. Mutations of Ret/ptc2 which fail to interact with Enigma also lose their ability to stimulate mitogenic signaling. Co-expression of LIM domains of Enigma blocked Ret/ptc2 stimulated DNA synthesis, indicating that Enigma is involved in the mitogenic signaling of Ret. These studies have indicated that two LIM domains of Enigma can interact with two receptor tyrosine kinases through tyrosine-based motifs with specificity residing in both the LIM domains and in the target structures.

14. SUBJECT TERMS	15. NUMBER OF PAGES
Endocytosis, Tyrosine Kinase, Growth Factor Receptor, ErBb2, EGFR, Breast Cancer	49
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

FOREWORD

Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the US Army.

✓ Where copyrighted material is quoted, permission has been obtained to use such material.

✓ Where material from documents designated for limited distribution is quoted, permission has been obtained to use the material.

✓ Citations of commercial organizations and trade names in this report do not constitute an official Department of Army endorsement or approval of the products or services of these organizations.

___ In conducting research using animals, the investigator(s) adhered to the "Guide for the Care and Use of Laboratory Animals," prepared by the Committee on Care and Use of Laboratory Animals of the Institute of Laboratory Resources, National Research Council (NIH Publication No. 86-23, Revised 1985).

N/P For the protection of human subjects, the investigator(s) adhered to policies of applicable Federal Law 45 CFR 46.

✓ In conducting research utilizing recombinant DNA technology, the investigator(s) adhered to current guidelines promulgated by the National Institutes of Health.

✓ In the conduct of research utilizing recombinant DNA, the investigator(s) adhered to the NIH Guidelines for Research Involving Recombinant DNA Molecules.

✓ In the conduct of research involving hazardous organisms, the investigator(s) adhered to the CDC-NIH Guide for Biosafety in Microbiological and Biomedical Laboratories.

Gordon N Gill 7/11/96
PI - Signature Date

Rumph Wu 7/24/96

Table of Contents

Introduction

The transmission of signals from outside of the cell to intracellular compartments is necessary in all cellular systems. Receptor tyrosine kinases function as an important group of molecules which propagate environmental signals to intracellular compartments by coupling to specific signal transduction pathways. Receptor tyrosine kinases are essential for normal growth, development and differentiation (1).

The major attenuation mechanism for signaling of cell surface receptors involves ligand-induced internalization and downregulation of receptors, a process dependent on both intrinsic tyrosine kinase activity and specific sequence motifs, termed endocytic codes. When receptors are excessively expressed, the endocytic mechanism is saturated so that attenuation can not occur and excessive signaling ensues from the cell surface.

The function of membrane receptors is dependent on correct trafficking between cellular compartment. The low density lipoprotein (LDL) and transferrin (Tf) receptors, whose primary function is transport of nutritional molecules into cells, are internalized by virtue of their constitutive binding to clathrin coated pits (2). Endocytosis of this class of receptors is independent of ligand binding but depends on specific cytoplasmic motifs that contain an aromatic residue residing in a tight turn structure (3,4). Receptor tyrosine kinases that activate signal transduction pathways display a different pattern of internalization. Ligand binding activates their intrinsic tyrosine kinase activity and results in redistribution of diffuse cell surface receptors to coated pits, enhanced rates of endocytosis and downregulation from the surface (5,6). Occupancy-induced internalization of epidermal growth factor (EGF) and insulin (Ins) receptors requires both intrinsic protein tyrosine kinase activity and specific sequence motifs. The juxtamembrane domain encoded by exon16 of InsR gene contains sequences (Gly-Pro-Leu-Tyr, to a lesser extent, Asn-Pro-Gln-Tyr) that are necessary for normal endocytosis. InsR deleted or mutated in these sequences retain tyrosine kinase activity but fail to undergo endocytosis (7). Unlike the LDLR and TfR which contain a short sequence of cytoplasmic residues and lack a tyrosine kinase domain (3,4), both tyrosine kinase activity and exon16 of InsR are required for endocytosis of InsR. Although a Tyr residue within the endocytic code is essential for internalization of InsR, this Tyr residue is not an auto-phosphorylation site. Tyrosine kinase activity is thought to lead to auto-phosphorylation and a conformational change that exposes the buried endocytic sequences in the normally dimerized InsR.

With identification of the endocytic codes has come an opportunity to use molecular recognition as a criterion for isolating specific proteins and subsequently defining the complex machinery involved in receptor trafficking. The two hybrid system, which depends upon molecular recognition properties of proteins for their isolation, was used as an approach to

identify a specific recognition protein for the endocytic codes of InsR. Enigma was isolated by interaction with Exon 16 of InsR (8) and was shown to recognizes the endocytic codes of InsR proportional to the strength of the codes in holo-InsR in vivo (7,8). Enigma interacts with the endocytic codes of InsR through one of the three LIM domains. The ability to specifically recognize the endocytic code of InsR via a LIM domain fulfills the first property of the endocytosis mechanism.

In addition to InsR, Enigma specifically interacted with the receptor tyrosine kinase Ret (9). The Ret oncogene was originally identified as a transforming gene by transfection of T cell lymphoma DNA into NIH3T3 cells (10). This oncogene was not expressed in the original lymphoma, but instead had arisen from a rearrangement during the transfection, hence the name Ret. The Ret proto oncogene encodes a member of the receptor tyrosine kinase gene superfamily with an extracelluler ligand-binding domain, a transmembrane domain and an intracellular domain (11,12).

The mutations of the Ret proto oncogene have been linked to four human disease syndromes: the dominantly inherited cancer syndromes FMTC, MEN2A and MEN2B, and Hirschsprung's disease (13-15). Deletions and nonsense point mutations leading to truncation of the Ret receptor, and single amino acid substitutions in the kinase domain of Ret have been reported in those patients. The Ret/ptc2 oncogene, an activated form of the Ret proto oncogene, has been found specifically in 11% to 15% of human papillary thyroid carcinomas (16). Ret/ptc2 is the product of a crossover between the genes coding for the tyrosine kinase domain of Ret and the type Iα regulatory subunit (RIα) of cAMP-dependent protein kinase (17,18). Both oncogenic forms of Ret are constitutively phosphorylated on tyrosine, display an in vitro auto-phosphorylation activity, and are translocated from the membrane to cytoplasm.

The mitogenic activity of Ret/ptc2 has been studied by a microinjection-based assay . RIα and a tyrosine residue located carboxyl terminal to the kinase core of Ret are required for the mitogenesis of Ret/ptc2 (19). Although the SH2 domains of PLC-γ and Grb10 binds to Ret/ptc2, these interactions are not required for mitogenic signaling (20). Ret/ptc2 also binds to one of the three LIM domains of Enigma (9). In contrast to the SH2 interactions, disruption of the LIM-Ret interaction abolishes Ret/ptc2 mitogenic signaling (20). This suggests that Ret/ptc2 signal transduction is mediated by LIM domain recognition of a Tyr-base motif but not a common signaling molecule the SH2 domain. This is the first description of a LIM domain protein involved in mitogenic signaling.

Given that Enigma, which contains three LIM domains at its carboxyl terminus, was found to interact with two receptor tyrosine kinases, InsR and Ret, determining the molecular basis of LIM domains recognition becomes important.

The LIM motif was first identified in three developmentally important transcription factors, C. elegans **L**in-11, rat **I**sl-1 and C. elegans **M**ec-3, from which the acronym LIM was derived in 1990 (21-23). The LIM domain is a Cys-rich sequence motif that is found in a variety of proteins with diverse functions and subcellular distributions, including transcription factors, proto-oncogene products, and components of adhesion plaques and the actin-based cytoskeleton (24,25). The LIM motif displays the Cys-rich consensus sequence $CX_2CX_{16-23}HX_2CX_2CX_2CX_{16-21}CX_{2-3}$(C, D, H). It has been demonstrated by spectroscopic analysis that the LIM domain defines a specific zinc-binding structure (26).

Structural analysis by NMR of the carboxyl terminal LIM domain of CRP showed that the second zinc-binding module has a structure similar to the zinc fingers in the GATA-1 transcription factor and the glucocorticoid receptor (27). Despite this structural similarity, there is no experimental evidence for a DNA binding function for LIM domains. However, there is strong evidence for their role in protein•protein interaction. The LIM-only protein CRP can form heterodimers with the complex LIM protein Zyxin (28). Interaction with CRP is specific for LIM1 of Zyxin and neither LIM2 nor LIM 3 of Zyxin nor LIMs of Mec-3 bind. CRP can also form homodimers through its LIM domains but no specificity of LIM•LIM interaction is observed (29). Binding of a LIM domain to a very different protein domains has been shown by the specific association of RBTN1, RBTN2 with the helix-loop-helix domains of TAL1, TAL2 and LYL1 (30,31). LIM domains in the transcription factor Lmx-1 are necessary for specific functional interaction with the helix-loop-helix protein shPan-1 which lacks LIM domains (32), supporting the idea that LIM domains interact with structural motif distinct from other LIM domains.

Enigma which contains three LIM domains at its carboxyl terminus was found to interact with two receptor tyrosine kinases. LIM2 of Enigma specifically recognizes Ret whereas LIM3 of Enigma specifically recognizes InsR (8,9). Both LIM2 and LIM3 recognize Tyr-containing motifs located outside the tyrosine kinase cores of Ret and InsR. The ability of LIM3 of Enigma to recognize the active endocytic codes of InsR fulfills the first property of the endocytic mechanism. Overexpression of LIM domains of Enigma blocked the mitogenic signaling of Ret/ptc2, indicating that Enigma is involved in the mitogenic signaling of Ret/ptc2 (20). These observations indicate LIM domains of Enigma function as protein interaction modules with specificity provided by both the LIM domains and target structures. Enigma is a potentially interesting protein that is involved in signal transduction and may also play a role in endocytosis of receptor tyrosine kinase.

Results

1. Isolation and characterization of Enigma

The observation that occupancy-induced high-affinity saturable endocytosis depends upon specific 4-6 amino acid sequence codes in signaling receptors predicts a specific cellular "internalization component" that specifically recognizes the code and functions in this process. The yeast two-hybrid system described by Zervos et al (33,34) was used to isolate a putative endocytic code recognition protein.

Reagents for the yeast two-hybrid system were obtained from Dr. Roger Brent (34). Exon 16 of InsR was subcloned into the LexA-fusion vector pEG202 and coexpressed in the EGY48 strain of Saccharomyces cerevisiae with an oligo(dT) -prepared HeLa cDNA library. From approximately one million co-transformants, we isolated 398 Leu^+ colonies. Three of these were blue only on galactose X-gal but not on glucose X-gal medium, confirming that they require the library encoded protein. The library plasmids were rescued from yeast cells and shown to be identical. Sequence of the cDNA fragment was used to search the data bank and shown to be unique. This novel protein, based on its code recognition properties, is termed Enigma (8).

The cDNA fragment was used as a hybridization probe to isolate a full-length cDNA from a λgt11 cDNA library prepared from SK-N-MC cells (Fig. 1). The cDNA correpsonds to a single mRNA species of 1.7 kilobases determined by Northern blotting. The cDNA contains a Kozak sequence at an initiator Met, and an open reading frame encoding a 455 amino acid protein. In vitro translation yielded a 55-kDa protein as expected from the open reading frame. Western blotting with an affinity-purified antibody generated against Enigma confirmed a single cytoplasmic ~55-kDa protein in the HeLa cells in agreement with the open reading frame and in vitro translation.

Although Enigma is a novel protein not present in the data banks, Enigma contains three LIM domains within its carboxyl terminus. The interacting domain isolated from yeast consisted predominantly of one LIM domain, the carboxyl terminal LIM domain (LIM3).

LIM3 of Enigma recognized exon 16 of InsR in the two hybrid system. To further examine the interaction in another approach, the LIM3 of Enigma was expressed as GST fusion protein and was tested for its ability to bind holo InsR. Because ligand-induced internalization of InsR requires intrinsic tyrosine kinase activity and sequence encoded by exon 16, we also asked whether tyrosine kinase activity is required for the association between Enigma and InsR. Insulin activated the InsR tyrosine kinase activity and stimulated autophosphorylation of InsR (36). The GST-LIM3 of Enigma and GST protein were immobilized on glutathione agarose beads and mixed with partially purified InsR without or with 100nM of insulin treatment. The GST-LIM3 of Enigma precipitated the InsR regardless

of whether insulin was added while the GST protein did not bind either form of InsR. Treatment with insulin activated the tyrosine kinase activity and auto-phosphorylation of InsR but did not affect the ability of InsR to bind GST-LIM3 of Enigma. These data demonstrate that LIM3 of Enigma recognizes holo InsR as well as exon 16 independently of Tyr phosphorylation of InsR.

For the endocytosis of InsR, tyrosine kinase activity of InsR is essential and necessary. The interaction of Enigma with InsR independently of tyrosine kinase activity of InsR indicates that the Enigma-InsR complex may be necessary but is not sufficient to support the internalization of InsR. Additional components which are regulated by the tyrosine kinase activity of InsR are also required.

2. Specific Interaction of Enigma with endocytic codes

Exon 16, which encodes 22 amino acids of the submembrane domain of the β-subunit of InsR, contains two endocytic codes: a strong code, Gly-Pro-Leu-Tyr, and a weak code, Asn-Pro-Glu-Tyr. Both codes form Tyr-containing tight turns in solution (37). Mutations of Gly-Pro-Leu-Tyr to Ala-Pro-Leu-Ala and Asn-Pro-Glu-Tyr to Ala-Pro-Glu-Ala reduced InsR internalization to 32% and 87% respectively (7).

To determine whether Enigma recognizes the active endocytic codes of InsR, the two hybrid system was used to measure the interaction of Enigma with mutant exon 16 sequence. Both the LIM3 and holo Enigma were used. Mutations of InsR exon 16 were cloned into the bait plasmid individually and interactions were determined by monitoring β-galactosidase activity. A 2-amino acid change in the strong endocytic code (Gly-Pro-Leu-Tyr to Ala-Pro-Leu-Ala) completely abolished the interaction. A 2-amino acid change in the weaker endocytic code (Asn-Pro-Glu-Tyr to Ala-Pro-Glu-Ala) decreased interaction but did not abolish it. No interaction occurred when mutation in both sequences were present. Enigma thus recognizes the strong code Gly-Pro-Leu-Tyr, and to a lesser extent, the weak code Asn-Pro-Glu-Tyr in exon 16 of InsR (Fig. 2). There is thus strong concordance between effects of these point mutants in InsR in vivo and in vitro interaction between Enigma and exon 16.

Although many receptors internalize via coated pits, specificity is implied by observations that EGFR do not compete with TfR for internalization, nor do InsR compete with mannose 6-phosphate receptors for internalization (6). To investigate specificity, interaction of LIM3 and holo Enigma with endocytic codes from other receptors was also studied by the two-hybrid system. Small fragments containing the individual endocytic codes of EGFR, IGF-1R, LDLR and TfR were cloned in the pEG202 plasmid, expressed as LexA fusion proteins and analyzed for their ability to interact with Enigma (Fig. 2). Neither LIM3 nor full length Enigma recognized the endocytic codes in IGF1-R, EGFR, LDLR and TfR. Interaction of LIM3 of Enigma with the Gly-Pro-Leu-Tyr strong code of InsR thus shows high specificity.

A synthetic sequence Asn-Asn-Ala-Tyr-Phe, predicted to have structural features of a tight turn, effectively replaced the Gln-Gln-Gly-Phe-Phe endocytic code for EGFR internalization (38). When Asn-Asn-Ala-Tyr-Phe was substituted for a single endocytic code of EGFR, it was not recognized by Enigma. A significant interaction was observed when Asn-Asn-Ala-Tyr-Phe was expressed in two copies in a predicted helix-turn-helix as occurs in native EGFR where two endocytic codes are separated by a predicted helical region. The interaction of full length Enigma was stronger than the interaction of LIM3 with Asn-Asn-Ala-Tyr-Phe.

The structural features of LIM domains are highly conserved. To determine whether LIM domains have target specificity, individual LIM domains of Enigma and LIM domains from other proteins were examined for their ability to interact with InsR endocytic codes. LIM domains of Enigma, CRP, Zyxin, Mec-3, Isl-1, Lmx -1 and Paxillin were cloned into pJG4-5 and their ability to associate with endocytic codes of EGFR, IGF-1R, LDLR and TfR were analyzed. Similar to the results seen with full length Enigma, most LIM domains failed to recognize endocytic codes of IGF-1R, LDLR, TfR, and EGFR. Only LIM3 of Enigma recognized the endocytic codes of InsR.

3. Identification of the recognition motif for LIM3 of Enigma

To determine the recognition motif for the LIM3 domain of Enigma, a random peptide library selection technique was used to study the consensus binding sequence for LIM3 of Enigma. Random peptide library selection has been successfully used to determine the sequence specificity of the peptide-binding sites of SH2 and SH3 domains as well as the optimal substrates of protein kinases (39,40).

We constructed a fixed Tyr peptide library comprising peptides of sequence: **Met-Ala-**X-X-X-X-**Tyr-**X-X-X-X-**Ala-Lys-Lys-lys**, where X indicates all amino acid except Trp, Cys, Ser, Thr or Tyr. LIM3 of Enigma was expressed as a GST fusion protein, immobilized on glutathione agarose and incubated with the Tyr peptide library. Unbound peptides were washed away and bound peptides were released by acid and subjected to micro-sequencing. The amino acids preferentially selected by LIM3 of Enigma at positions -4, -3, -2 and -1 amino terminal to the Tyr residue and +1, +2, +3 and +4 carboxyl terminal to the Tyr were shown in Fig. 3A. The greatest selectivity was observed at positions -1 and +2 where Pro was preferred. At the -2 position, Gly was the preferred amino acid and at the +3 position, both Val and Ile were highly selected. Phe was preferred at position +4.

Since Pro was preferred at positions -1 and +2, a second library with the sequence Met-Ala-X-X-X-X-X-Pro-X-X-Pro-X-X-X-Ala-Lys-Lys-Lys in which Pro was fixed with two intervening amino acids was designed to further test selectivity. This library also included Tyr at the ten degenerate positions. The general motif determined by this library was similar to

those found with the Tyr fixed peptide library (Fig. 3B). In addition, a Tyr residue was highly selected at position +4. By comparison of these two motifs from two peptide library selections, the peptide sequence of Gly-Pro-Hyr-Gly-Pro-Hyr-Tyr/Phe-Ala was determined to be the recognition motif for LIM3 of Enigma (Fig. 3C). This peptide sequence is highly homologous to the sequence of exon16 of InsR.

To confirm the binding motif for LIM3 of Enigma, peptides were used to disrupt the complex of LIM3/Enigma with holo InsR. HIRC cell lysate containing InsR was incubated with GST-LIM3 of Enigma without or with competitor peptides. Binding of InsR to GST-LIM3 of Enigma was also inhibited by a 12 amino acid peptide (Asp-Gly-Pro-Leu-Gly-Pro-Leu-Tyr-Ala-Ser-Ser-Asn) corresponding to exon 16 of InsR but not by two mutant peptides (Asp-Gly-Pro-Leu-**Ala**-Pro-Leu-**Ala**-Ala-Ser-Ser-Asn and Asp-Gly-**Ala**-Leu-Gly-**Ala**-Leu-Tyr-Ala-Ser-Ser-Asn). These data confirmed two Pro and one Tyr residues were essential for the recognition of LIM3 of Enigma determined by the random peptide library selection. One amino acid substitution of Leu (the residue before Tyr) for Ile did not effect its ability to compete for InsR binding, confirming the random peptide library selection of a hydrophobic residue at position +3. These peptide competition results demonstrate that both the Pro and Tyr residues are required to mediate interaction of LIM3 of Enigma with exon 16 of InsR.

4. Differential recognition of Ret/ptc2 and InsR by LIM domains of Enigma

When an oncogenic form of Ret, Ret/ptc2, was used in a yeast two hybrid system to identify interacting proteins, several SH2 and SH3 domain containing proteins and Enigma were isolated (20). To determine the domains of Enigma responsible for the interaction with Ret/ptc2, regions of Enigma were cloned into pJG4-5 and tested for their ability to bind Ret/ptc2 expressed as a LexA fusion protein in pEG202 in yeast. As shown in Fig. 4, full length Enigma bound to Ret/ptc2 and exon16 of InsR to a similar extent. The amino terminal 279 amino acids of Enigma did not interact with either Ret/ptc2 or exon16 of InsR. The carboxyl terminal 275 amino acid containing the three LIM domains were responsible for protein-protein associations with both InsR and Ret. When the carboxyl terminus was further divided into individual LIM domains, LIM2 of Enigma bound Ret/ptc2 but not exon 16 of InsR. Conversely, LIM3 of Enigma bound exon16 of InsR but not Ret/ptc2. These results demonstrated that LIM2 of Enigma was both necessary and sufficient to support the association of Enigma with Ret/ptc2 and could be physically separated from LIM3 of Enigma which was responsible for Enigma association with InsR. LIM1 of Enigma, which bound two atoms of zinc characteristic of LIM domains (personal communication from Dr. Dennis Winge), did not associate with either Ret/ptc2 or InsR.

The specificity of interaction of Ret/ptc2 and InsR with other LIM domains was further examined. As shown in Fig. 5, LIM domains of Mec-3, Isl-1, Lmx-1, Zyxin, CRP, or

Paxillin did not recognize Ret/ptc2 or exon 16 of InsR. The specificity of recognition thus resides in the LIM domains of Enigma.

To investigate whether tyrosine kinase activity was required for the association between LIM2/Enigma and Ret, the EGFR/Ret chimeric protein was used. Because the ligand for Ret tyrosine kinase receptor was unknown at the time of these studies, the chimera generated by fusing the extracellular and transmembrane domains of EGFR and the intracellular domain of Ret was used. EGF activated the Ret tyrosine kinase activity and mitogenic response of this chimera (41). 3T3 cells expressing the EGFR/Ret were treated without or with EGF and cell lysate was mixed with GST or the GST-LIM domains of Enigma. GST-LIM2 of Enigma interacted with phosphorylated as well as unphosphorylated EGFR/Ret receptors. The interaction of LIM2 of Enigma with Ret is thus independent of ligand activation and Ret autophosphorylation.

5. Mapping of the site of interaction of Ret with Enigma.

Mitogenic activity of Ret/ptc2 requires dimerization via RIα and the Tyr^{586} residue located carboxyl terminal to the kinase core of Ret (19). To determine the region of Ret responsible for association with LIM2 of Enigma, Ret/ptc2 was divided into RIα and Ret tyrosine kinase domains (Ret/tk) and their ability to bind LIM domains of Enigma was tested in yeast. Ret/ptc2 and mutants were expressed as LexA fusion proteins in pEG202 and the three LIM domains of Enigma were cloned into the pJG4-5 vector. As shown in Fig. 6, Ret/tk recognized the LIM domains of Enigma but the RIα fragment did not. Deletion of the carboxyl terminal 23 amino acid of Ret (C'574) distal to the conserved tyrosine kinase core abolished Enigma binding, demonstrating that the carboxyl terminal 23 amino acids of Ret was necessary for Ret binding to Enigma.

These results indicate the carboxyl terminal 23 amino acids of RET are necessary for Ret binding to Enigma. To determine whether the carboxyl terminal region of Ret is sufficient to support the binding, the carboxyl terminal 61 amino acids of Ret (residues 536 to 596) were expressed as a GST fusion protein (GST-C'/Ret) and tested for Enigma binding. GST and GST-C'/Ret were immobilized on glutathione agarose and mixed with full length, the amino terminal (N279) or the carboxyl terminal (C275) domains of Enigma that were expressed as HA epitope-tagged fusion proteins in 293 cells (Fig. 7). Equal amounts of GST and GST-C'/Ret were assessed for their ability to bind Enigma proteins. GST-C'/Ret interacted with full length Enigma as well as with the carboxyl terminus containing the three LIM domains of Enigma but failed to bind the amino terminus of Enigma. There was no binding of any of the three forms of Enigma protein to GST. Theses results demonstrate that the carboxyl terminal 61 amino acids of Ret determine the association of Ret with Enigma.

6. Requirement of Enigma for mitogenic signaling of Ret/ptc2

To examine the biological significance of the interaction between Ret/ptc2 and Enigma, the ability of various Ret/ptc2 constructs to stimulate mitogenic signaling was compared with their ability to interact with Enigma. Mutations of Tyr to Phe at residues 429 and 539 did not effect Ret/ptc2 binding to Enigma while mutation of Tyr^{586} to Phe eliminated Ret/ptc2 interaction with Enigma (Fig. 8). Tyr^{429} and Tyr^{539} of Ret/ptc2 were identified as the binding sites for Grb10 and PLC-γ respectively (20). Mutations of Ret/ptc2 which failed to associate with Grb10 and PLC-γ had no significant effect on mitogenic activity, while mutation of Tyr^{586} to Phe (Y586F) or deletion of the carboxyl terminal 23 amino acids (C'574) of Ret/ptc2 completely abolished the ability to stimulate DNA synthesis (Fig. 9). The strong correlation between mutant that disrupt Enigma binding and loss of mitogenic activity suggested that Enigma was required for the mitogenic signaling of Ret/ptc2.

To further detremine the functional role of Enigma in the mitogenic signaling of Ret/ptc2, co-injection experiments were performed to attempt to block mitogenesis of Ret/ptc2. As shown in Fig. 10, co-injection of Ret/ptc2 with a plasmid which expressed the three LIM domains of Enigma (C275) blocked Ret/ptc2-induced DNA synthesis, while co-injection with a full length Enigma expression construct (FL) had no effect. These results support the conclusion that Ret/ptc2 requires Enigma for mitogenic signaling. If Enigma was simply competing for binding with some other signaling protein then both full-length Enigma and the LIM domains alone should block signaling. Additionally, the amino terminus of Enigma lacking LIM domains (N279) did not block Ret/ptc2 mitogenic signaling. The mitogenic block was specific for the LIM domains of Enigma because the LIM domains of Zyxin were without effect. It was also specific for Ret since the LIM domains of Enigma did not bind EGFR or block the mitogenic activity of the EGFR tyrosine kinase analog of Ret/ptc2 (RI/EGFR) previously shown to have mitogenic activity (9).

These studies indicate Enigma is necessary for Ret/ptc2 mitogenic signaling, but Ret kinase activity is also required (9). This tyrosine kinase activity is not required for recruitment of Enigma to Ret/ptc2. Additional components which are regulated by tyrosine kinase activity of Ret must also be required.

Conclusion

1. Enigma and InsR endocytic codes

The ability to specifically recognize the active endocytic code of InsR via a LIM domain fulfills the first property of the endocytic mechanism predicted from study of kinetics of ligand-induced saturable high affinity endocytosis of InsR, that of recognition, but additional criteria are necessary to demonstrate Enigma function in this process. Enigma recognized the endocytic codes of exon16 of InsR proportional to the strength of the codes in InsR in vivo (7,8), but failed to recognized other endocytic codes. Specificity resided not only in the tyrosine-containing tight turn but in the LIM domain as shown by the inability of LIM1 or 2 of Enigma, LIM domains of Zyxin, CRP, Isl-1, Lmx-1, Mec-3 and Paxillin to interact with exon 16. Although these observation indicate specificity for both partners to the interaction, they do not prove a function for Enigma which may act in ways unrelated to insulin action. The two hybrid system has high sensitivity for detecting protein interactions, and it remains possible that physiologically relevant interactions with exon 16 and with Enigma will be quite different from those described. The interaction between exon 16 of InsR and LIM3 of Enigma does nonetheless provide evidence for a novel protein interaction mechanism.

2. Enigma and Ret mitogenic signaling

In contrast to LIM3 of Enigma binding to InsR, LIM2 of Enigma associated specifically with Ret. LIM2 of Enigma interacted with the carboxyl terminus of Ret which is essential for Ret mitogenic signaling. Similar to InsR binding to LIM3 of Enigma, Ret was specific for LIM2 of Enigma and did not bind other LIM domains tested. The specificity of interaction was also demonstrated by the inability of LIM domains of Enigma to interact with either full length or the intracellular domain of EGFR. In a nuclear microinjection assay, the mitogenic activity of Ret/ptc2 was abolished by the carboxyl terminal truncation to residue 574 or by the mutation of Y^{586} to Phe. Co-expression of three LIM domains of Enigma blocked the mitogenic activity of Ret/ptc2 but did not effect the ability of a construct analogous to Ret/ptc2 encoding RIα residues 1-236 fused to the EGFR intracellular domain residues 647-1186 to stimulate DNA synthesis. These observation indicated that Enigma is involved in the mitogenic signaling of Ret/ptc2 but not of EGFR. The amino terminal portion of Enigma lacking LIM domains is required for mitogenic signaling because the LIM domains alone ablated the mitogenic signal. This is the first description of a tyrosine kinase oncogene that requires a LIM domain containing protein for mitogegic signaling.

3. Role of tyrosine kinase activity

LIM3 of Enigma recognized the major endocytic code of InsR while LIM2 of Enigma interacted with the carboxyl terminus of Ret which is essential for Ret mitogenic signaling. Interaction of two LIM domains of Enigma with these receptors did not require either tyrosine

kinase activity of these two receptors or tyrosine auto-phosphorylation on their target sequences. However a Tyr residue is the important structural feature for both interactions. Mutation of Tyr^{586} to Phe of Ret/ptc2 dramatically decreased its ability to interact with Enigma while mutation of Tyr^{953} to Ala of InsR abolished Enigma binding. Unphosphorylated or kinase-inactive Ret and InsR had similar affinities for association with Enigma, indicating the molecular structure of a Tyr residue rather than phosphorylation of the Tyr residue is the recognition feature between Enigma and these two receptor tyrosine kinases. For both the endocytosis of InsR and the mitogenic signaling of Ret, tyrosine kinase activity of InsR and Ret is essential and necessary. This indicates that the interaction of Enigma with InsR and Ret is essential but not sufficient to support biological processes. Additional components which are regulated by tyrosine kinase activity of these two receptors are also required.

4. Specificity of LIM domain recognition

LIM2 of Enigma specifically interacts with Ret while LIM3 of Enigma associates with InsR. The Asn-Lys-Leu-Tyr sequence located at the carboxyl terminus of Ret is essential for formation of the Ret and Enigma complex. For the interaction of InsR and Enigma, the Gly-Pro-Leu-Gly-Pro-Leu-Tyr sequence of the juxtamembrane region of InsR is required. Both LIM2 and LIM3 recognize Tyr containing motifs outside of the tyrosine kinase cores of Ret and InsR. Although the recognition motifs for LIM2 of Enigma and for LIM3 of Enigma share some structural similarity, they are not exchangeable. LIM2 of Enigma is specific for Ret and LIM3 of Enigma only interacts with InsR, demonstrating that these two LIM domains have the ability to distinguish between two Tyr-based motifs. Individual LIM domains within a single protein can clearly display distinct partner preferences. Because the structural features of the LIM domains are highly conserved, sequences other than the conserved residues that are involved in metal binding must be important for defining the selectivity of individual LIM domains for their particular partner.

Asn-Asn-Ala-Tyr-Phe was originally chosen to mimic or change the native EGFR sequence Gln-Gln-Gly-Phe-Phe. In the context of mutant EGFR it functions as an endocytic code and is the only sequence found that restored endocytosis to kinase-inactive EGFR. Although the structure of Asn-Asn-Ala-Tyr-Phe is unknown, its placement between sequences corresponding to helices in protein kinase A was expected to provide exposure perhaps as a helix-turn-helix structure. The interaction of several LIM domains with this structure provides support for the hypothesis that LIM domains recognize exposed Tyr-containing motifs (8).

5. Function of the LIM domains

An increasing number of proteins have been identified to display one or more copies of LIM domains. The LIM domain-containing proteins function in a variety of pathways and locations within the cell, implicating LIM domains as versatile peptide sequences that are

capable of acting in diverse cellular contexts. Although the NMR structure of LIM2 of CRP resembles the DNA binding domain of the GATA-1 transcription factor, no direct evidence that a LIM domain interacts with nucleic acid has been presented. Indeed, a lack of affinity for target DNA sequence has been reported for the LIM domains of Mec-3 (43).

Most available evidence has indicated that LIM domains function in specific protein-protein interactions. Like SH2 and SH3 domains, LIM domains are distributed in a variety of proteins and are proposed to function in assembly and disassembly of protein complexes. Interaction between two LIM domain-containing proteins, Zyxin and CRP, suggested that self-association may prove important in addition to recognition of motifs such as tight turns (8,28). LIM domains in the transcription factor Lmx-1 are necessary for specific functional interaction with shPan-1 which lacks LIM domains, supporting the idea that LIM domains interact with structural motifs distinct from other LIM domains (32).

SH2 and SH3 domains, although widely distributed in proteins, exhibit specificity in interaction, suggesting this must be true of LIM domains to account for their wide distribution in proteins of differing function. Multiple copies of each component, multiple Tyr-containing structures within one protein and multiple LIM domains in another protein, could increase the strength of interaction and broaden recognition properties.

A single LIM domain has been shown to serve as a specific interface for protein-protein interaction. Enigma contains three LIM domains within its carboxyl terminus and two of three LIM domains have recognized target motifs. LIM1 of Enigma that failed to recognize either Ret or InsR is likely to have a yet unidentified target protein in cells. Two LIM domains in a single protein can interact with their respective target proteins, suggesting that proteins with multiple LIM domains could function as a adapter molecules in assembly of protein complexes. Alternatively, multiple LIM domains could function to differentially assemble proteins with distinct receptors. The finding that binding of InsR and of Ret to Enigma does not require the protein tyrosine kinase activity of either receptor distinguishes LIM domain interaction from those of SH2 and PTB domains which direct assembly that is dependent on tyrosine kinase activity and covalent modifications of proteins (39,44,45). Enigma might become phosphorylated upon Ret activation or serve to localize Ret to a position required for kinase-mediated signaling.

Studies on LIM domains indicate that LIM domains functions in protein-protein interaction. The LIM domains belong to a collection of protein-binding interfaces that including SH2 and SH3 domains (39), PTB domain (44,45) and the helix-loop-helix motif (46). Whether the LIM domain participates in homotypic or heterotypic association with other proteins, these interactive domains may function to localize a protein to a particular subcellular compartment, to mediate the assembly of a multimeric protein machine, or more generally to

regulate the activity of partner proteins. The identification of interacting partners for LIM domains would provide an insight into pathways important for signal transduction, development, growth and cytoskeletal rearrangements.

References

1. van der Geer, P., Hunter, T. and Lindberg, R.A. (1994) Receptor protein-tyrosine kinases and their signal transduction pathways. *Annu Rev Cell Biol* **10**, 251-337
2. Goldstein, J.L., Brown, M.S., Anderson, R.G., Russell, D.W. and Schneider, W.J. (1985) Receptor-mediated endocytosis: concepts emerging from the LDL receptor system. *Annu Rev Cell Biol* **1**, 1-39
3. Bansal, A. and Gierasch, L.M. (1991) The NPXY internalization signal of the LDL receptor adopts a reverse-turn conformation. *Cell* **67**, 1195-1201
4. Collawn, J.F., Stangel, M., Kuhn, L.A., Esekogwu, V., Jing, S.Q., Trowbridge, I.S. and Tainer, J.A. (1990) Transferrin receptor internalization sequence YXRF implicates a tight turn as the structural recognition motif for endocytosis. *Cell* **63**, 1061-1072
5. Chen, W.S., Lazar, C.S., Poenie, M., Tsien, R.Y., Gill, G.N. and Rosenfeld, M.G. (1987) Requirement for intrinsic protein tyrosine kinase in the immediate and late actions of the EGF receptor. *Nature* **328**, 820-823
6. Backer, J.M., Shoelson, S.E., Haring, E. and White, M.F. (1991) Insulin receptors internalize by a rapid, saturable pathway requiring receptor autophosphorylation and an intact juxtamembrane region. *J Cell Biol* **115**, 1535-1545
7. Rajagopalan, M., Neidigh, J.L. and McClain, D.A. (1991) Amino acid sequences Gly-Pro-Leu-Tyr and Asn-Pro-Glu-Tyr in the submembranous domain of the insulin receptor are required for normal endocytosis. *J Biol Chem* **266**, 23068-23073
8. Wu, R.Y. and Gill, G.N. (1994) LIM domain recognition of a tyrosine-containing tight turn. *J Biol Chem* **269**, 25085-25090
9. Wu, R., Durick, K., Songyang, Z, Cantlet, L. C., Taylor, S. S. and Gill, G. N. (1996) Specificity of LIM domains interaction with receptor tyrosine kinases. *J Biol Chem* **271**, 15934-15942
10. Takahashi, M., Ritz, J. and Cooper, G.M. (1985) Activation of a novel human transforming gene, ret, by DNA rearrangement. *Cell* **42**, 581-588
11. Takahashi, M. and Cooper, G.M. (1987) ret transforming gene encodes a fusion protein homologous to tyrosine kinases. *Mol Cell Biol* **7**, 1378-1385
12. Takahashi, M., Buma, Y., Iwamoto, T., Inaguma, Y., Ikeda, H. and Hiai, H. (1988) Cloning and expression of the ret proto-oncogene encoding a tyrosine kinase with two potential transmembrane domains. *Oncogene* **3**, 571-578
13. Mulligan, L.M., Eng, C., Healey, C.S., Clayton, D., Kwok, J.B., Gardner, E., Ponder, M.A., Frilling, A., Jackson, C.E., Lehnert, H. and et, a. (1994) Specific mutations of the RET proto-oncogene are related to disease phenotype in MEN 2A and FMTC. *Nat Genet* **6**, 70-74
14. Hofstra, R.M., Landsvater, R.M., Ceccherini, I., Stulp, R.P., Stelwagen, T., Luo, Y., Pasini, B., Hoppener, J.W., van Amstel, H.K., Romeo, G., Lips, C. J. M. and Buys, C. H. C. M. (1994) A mutation in the RET proto-oncogene associated with multiple endocrine neoplasia type 2B and sporadic medullary thyroid carcinoma [see comments]. *Nature* **367**, 375-376
15. Romeo, G., Ronchetto, P., Luo, Y., Barone, V., Seri, M., Ceccherini, I., Pasini, B., Bocciardi, R., Lerone, M., Kaariainen, H. and Martucciello, G. (1994) Point mutations affecting the tyrosine kinase domain of the RET proto-oncogene in Hirschsprung's disease [see comments]. *Nature* **367**, 377-378
16. Grieco, M., Cerrato, A., Santoro, M., Fusco, A., Melillo, R.M. and Vecchio, G. (1994) Cloning and characterization of H4 (D10S170), a gene involved in RET rearrangements in vivo. *Oncogene* **9**, 2531-2535

17. Lanzi, C., Borrello, M.G., Bongarzone, I., Migliazza, A., Fusco, A., Grieco, M., Santoro, M., Gambetta, R.A., Zunino, F., Della Porta, G. and et, a. (1992) Identification of the product of two oncogenic rearranged forms of the RET proto-oncogene in papillary thyroid carcinomas. *Oncogene* **7**, 2189-2194
18. Bongarzone, I., Butti, M.G., Coronelli, S., Borrello, M.G., Santoro, M., Mondellini, P., Pilotti, S., Fusco, A., Della Porta, G. and Pierotti, M.A. (1994) Frequent activation of ret protooncogene by fusion with a new activating gene in papillary thyroid carcinomas. *Cancer Res* **54**, 2979-2985
19. Durick, K, Yao, V. J., Borrello, M. G., Bongarzone, I., Pierotti, M. A. S. and Taylor, S. S. (1994) Tyrosine outside the kinase core and dimerization are required for the mitogenic activity of RET/ptc2. *J Biol Chem* **270**, 24642-24645
20. Durick, K., Wu, R., Taylor, S. S., and Gill, G. N. (1996) Mitogenic signal by RET/ptc2 requires association with Enigma via a LIM domain. *J Biol Chem* **271**, 12691-12694
21. Karlsson, O., Thor, S., Norberg, T., Ohlsson, H. and Edlund, T. (1990) Insulin gene enhancer binding protein Isl-1 is a member of a novel class of proteins containing both a homeo- and a Cys-His domain. *Nature* **344**, 879-882
22. Freyd, G., Kim, S.K. and Horvitz, H.R. (1990) Novel cysteine-rich motif and homeodomain in the product of the Caenorhabditis elegans cell lineage gene lin-11. *Nature* **344**, 876-879
23. Way, J.C. and Chalfie, M. (1988) mec-3, a homeobox-containing gene that specifies differentiation of the touch receptor neurons in C. elegans. *Cell* **54**, 5-16
24. Dawid, I.B., Toyama, R. and Taira, M. (1995) LIM domain proteins. *C R Acad Sci Iii* **318**, 295-306
25. Gill, G. N. (1995) The enigma of LIM domains. *Structure* **3**, 1285-1289
26. Michelsen, J.W., Schmeichel, K.L., Beckerle, M.C. and Winge, D.R. (1993) The LIM motif defines a specific zinc-binding protein domain. *Proc Natl Acad Sci U S A* **90**, 4404-4408
27. Perez-Alvarado, G.C., Miles, C., Michelsen, J.W., Louis, H.A., Winge, D.R., Beckerle, M.C. and Summers, M.F. (1994) Structure of the carboxy-terminal LIM domain from the cysteine rich protein CRP [see comments]. *Nature Structural Biology* **1**, 388-398
28. Schmeichel, K.L. and Beckerle, M.C. (1994) The LIM domain is a modular protein-binding interface. *Cell* **79**, 211-219
29. Feuerstein, R., Wang, X., Song, D., Cooke, N.E. and Liebhaber, S.A. (1994) The LIM/double zinc-finger motif functions as a protein dimerization domain. *Proc Natl Acad Sci U S A* **91**, 10655-10659
30. Wadman, I., Li, J., Bash, R.O., Forster, A., Osada, H., Rabbitts, T.H. and Baer, R. (1994) Specific in vivo association between the bHLH and LIM proteins implicated in human T cell leukemia. *Embo J* **13**, 4831-4839
31. Valge-Archer, V.E., Osada, H., Warren, A.J., Forster, A., Li, J., Baer, R. and Rabbitts, T.H. (1994) The LIM protein RBTN2 and the basic helix-loop-helix protein TAL1 are present in a complex in erythroid cells. *Proc Natl Acad Sci U S A* **91**, 8617-8621
32. German, M.S., Wang, J., Chadwick, R.B. and Rutter, W.J. (1992) Synergistic activation of the insulin gene by a LIM-homeo domain protein and a basic helix-loop-helix protein: building a functional insulin minienhancer complex. *Genes Dev* **6**, 2165-2176
33 Zervos, A.S., Gyuris, J. and Brent, R. (1993) Mxi1, a protein that specifically interacts with Max to bind Myc-Max recognition sites. *Cell* **72**, 223-232
34. Gyuris, J., Golemis, E., Chertkov, H. and Brent, R. (1993) Cdi1, a human G1 and S phase protein phosphatase that associates with Cdk2. *Cell* **75**, 791-803
35. Turner, C.E. and Miller, J.T. (1994) Primary sequence of paxillin contains putative SH2 and SH3 domain binding motifs and multiple LIM domains: identification of a vinculin and pp125Fak-binding region. *J Cell Sci* **107**, 1583-1591
36. Olefsky, J.M. (1990) The insulin receptor. A multifunctional protein [published erratum appears in Diabetes 1991 Jan;40(1):159]. *Diabetes* **39**, 1009-1016

37. Backer, J.M., Shoelson, S.E., Weiss, M.A., Hua, Q.X., Cheatham, R.B., Haring, E., Cahill, D.C. and White, M.F. (1992) The insulin receptor juxtamembrane region contains two independent tyrosine/beta-turn internalization signals. *J Cell Biol* **118**, 831-839
38. Chang, C.P., Lazar, C.S., Walsh, B.J., Komuro, M., Collawn, J.F., Kuhn, L.A., Tainer, J.A., Trowbridge, I.S., Farquhar, M.G., Rosenfeld, M.G., Wiley, S. and Gill, G. N. (1993) Ligand-induced internalization of the epidermal growth factor receptor is mediated by multiple endocytic codes analogous to the tyrosine motif found in constitutively internalized receptors. *J Biol Chem* **268**, 19312-19320
39. Songyang, Z., Shoelson, S.E., Chaudhuri, M., Gish, G., Pawson, T., Haser, W.G., King, F., Roberts, T., Ratnofsky, S., Lechleider, R.J. and et, a. (1993) SH2 domains recognize specific phosphopeptide sequences. *Cell* **72**, 767-778
40. Songyang Z; Blechner S; Hoagland N; Hoekstra MF; Piwnica-Worms H; Cantley LC. (1994) Use of an oriented peptide library to determine the optimal substrates of protein kinases. *Current Biology* **4**, 973-982.
41. Santoro, M., Wong, W.T., Aroca, P., Santos, E., Matoskova, B., Grieco, M., Fusco, A. and di Fiore, P.P. (1994) An epidermal growth factor receptor/Ret chimera generates mitogenic and transforming signals: evidence for a ret-specific signaling pathway. *Mol Cell Biol* **14**, 663-675
42. Hildebrand, J.D., Schaller, M.D. and Parsons, J.T. (1993) Identification of sequences required for the efficient localization of the focal adhesion kinase, pp125FAK, to cellular focal adhesions. *J Cell Biol* **123**, 993-1005
43. Xue, D., Tu, Y. and Chalfie, M. (1993) Cooperative interactions between the Caenorhabditis elegans homeoproteins UNC-86 and MEC-3. *Science* **261**, 1324-1328
44. Pawson, T. (1995) Protein modules and signalling networks. *Nature* **373**, 573-580
45. Kavanaugh, W.M., Turck, C.W. and Williams, L.T. (1995) PTB domain binding to signaling proteins through a sequence motif containing phosphotyrosine. *Science* **268**, 1177-1179
46. Edmondson, D.G. and Olson, E.N. (1993) Helix-loop-helix proteins as regulators of muscle-specific transcription. *J Biol Chem* **268**, 755-758

Appendices

Enigma Gene Product

```
      GAATTCCCGTTGCTGTCGCCCAACGAGGCTCCCTGGAGCCGACGCAGAGCAGCGCCCTGGCCGGGCCAAGCAGGAGCCGGCATC  84

ATGGATTCCTTCAAAGTAGTGCTGGAGGGGCCAGCACCTTGGGGGTTCCGGCTGCAAGGGGGCAAGGACTTCAATGTGCCCCTCTCCATT 174
 M  D  S  F  K  V  V  L  E  G  P  A  P  W  G  F  R  L  Q  G  G  K  D  F  N  V  P  L  S  I  30

TCCCGGCTCACTCCTGGGGGCAAAGCGGCGCAGGCCGGAGTGGCCGTGGGTGACTGGGTGCTGAGCATCGATGGCGAGAATGCGGGTAGC 264
 S  R  L  T  P  G  G  K  A  A  Q  A  G  V  A  V  G  D  W  V  L  S  I  D  G  E  N  A  G  S  60

CTCACACACATCGAAGCTCAGAACAAGATCCGGGCCTGCGGGGAGCGCCTCAGCCTGGGCCTCAGCAGGGCCCAGCCGGTTCAGAGCAAA 354
 L  T  H  I  E  A  Q  N  K  I  R  A  C  G  E  R  L  S  L  G  L  S  R  A  Q  P  V  Q  S  K  90

CCGCAGAAGGCCTCCGCCCCCGCCGCGGACCCTCCGCGTACACCTTTGCACCCAGCGTCTCCCTCAACAAGACGGCCCAACCCTTTGGGC 444
 P  Q  K  A  S  A  P  A  A  D  P  P  R  T  P  L  H  P  A  S  P  S  T  R  R  P  N  P  L  G  120

CCCCCCGCCCCTGACAGCCCCCCGCAGCAGAATGGACAGCCGCTCCGACCGCTGGTCCCAGATGCCAGCAAGCAGCGGCTGATGGAGAAC 534
 P  P  A  P  D  S  P  P  Q  Q  N  G  Q  P  L  R  P  L  V  P  D  A  S  K  Q  R  L  M  E  N  150

ACAGAGGACTGGCGGCCGCGGCCGGGACAGGCCAGTCGCGTTCCTTCCGCATCCTTGCCCACCTCACAGGCACCGAGTTCATGCAAGACC 624
 T  E  D  W  R  P  R  P  G  Q  A  S  R  V  P  S  A  S  L  P  T  S  Q  A  P  S  S  C  K  T  180

CCGGATGAGGAGCACCTGAAGAAATCAAGCCAGGTGCCAGACAGAAGCCCCAGCCCCAGCCTCATCTACACCCCAGGAGCCCTGGCCTGG 714
 P  D  E  E  H  L  K  K  S  S  Q  V  P  D  R  S  P  S  P  S  L  I  Y  T  P  G  A  L  A  W  210

CCCTACCGCCCCCAGCCCTACCAGCCGCCCGCCCTGGGCTGTGGACCCTGCGTTTGCCGAGCGCTATGCCCCGGACAAAACGAGCACAGT 804
 P  Y  R  P  Q  P  Y  Q  P  P  A  L  G  C  G  P  C  V  C  R  A  L  C  P  G  Q  N  E  H  S  240

GCTGACCCACACAGCCAGCCAGCCACGCCCACGCCGCTGCAGAGCCGCACCTCCATTGTGCAGGCAGCTGCCGGAGGGGTGCCAGGAGGG 894
 A  D  P  H  S  Q  P  A  T  P  T  P  L  Q  S  R  T  S  I  V  Q  A  A  A  G  G  V  P  G  G  270

GGCAGCAACAACGGCAAGACTCCCGTGTGTCACCAGTGCCACAAGGTCATCCGGGGCCGCTACCTGGTGGCGCTGGGCCACGCGTACCAC 984
 G  S  N  N  G  K  T  P  V  C  H  Q  C  H  K  V  I  R  G  R  Y  L  V  A  L  G  H  A  Y  H  300

CCGGAGGAGTTTGTGTGTAGCCAGTGTGGGAAGGTCCTGGAAGAGGGTGGCTTCTTTGAGGAGAAGGGCGCCATCTTCTGCCCACCATGC 1074
 P  E  E  F  V  C  S  Q  C  G  K  V  L  E  E  G  G  F  F  E  E  K  G  A  I  F  C  P  P  C  330

TATGACGTGCCGCTATGCACCCAGCTGTGCCAAGTGCAAGAAGAAGATTACAGGCGAGATCATGCACGCCCTGAAGATGACCTGGCACGTG 1164
 Y  D  V  R  Y  A  P  S  C  A  K  C  K  K  K  I  T  G  E  I  M  H  A  L  K  M  T  W  H  V  360

CACTGCTTTACCTGTGCTGCCTGCAAGACGCCCATCCGGAACAGGGCCTTCTACATGGAGGAGGGCGTGCCCTATTGCGAGCGAGACTAT 1254
 H  C  F  T  C  A  A  C  K  T  P  I  R  N  R  A  F  Y  M  E  E  G  V  P  Y  C  E  R  D  Y  390

GAGAAGATGTTTGGCACGAAATGCCATGGCTGTGACTTCAAGATCGACGCTGGGGACCGCTTCCTGGAGGCCCTGGGCTTCAGCTGGCAT 1344
 E  K  M  F  G  T  K  C  H  G  C  D  F  K  I  D  A  G  D  R  F  L  E  A  L  G  F  S  W  H  420

GACACCTGCTTCGTCTGTGCGATATGTCAGATCAACCTGGAAGGAAAGACCTTCTACTCCAAGAAGGACAGGCCTCTCTGCAAGAGCCAT 1434
 D  T  C  F  V  C  A  I  C  Q  I  N  L  E  G  K  T  F  Y  S  K  K  D  R  P  L  C  K  S  H  450

GCCTTCTCTCATGTGtgagccccttctgcccacagctgccgcggtggcccctagcctgaggggcctggagtcgtggccctgcatttctgg 1524
 A  F  S  H  V 455

gtagggctggcaatggttgccttaaccctggctcctggcccgagcctgggctccctggccctgccccacccaccttatcctcccacccca 1614
ctccctccaccaccacagcacaccggtgctggccacaccagccccctttcacctccagtgccacaataaacctgtacccagctgaaaaaa 1704
aaaaaaaaaaaaaaactcgag 1725
```

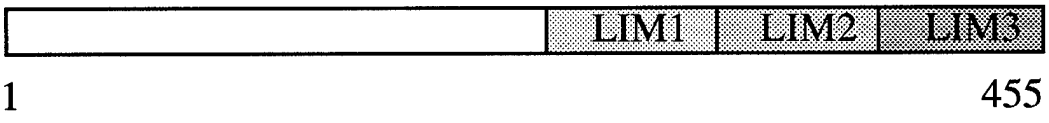


Fig. 1 **The nucleotide and amino acid sequence of Enigma.**
The carboxyl terminus of Enigma isolated from a HeLa cDNA library using the two-hybrid system to detect interaction with the 22 amino acids encoded by exon 16 of InsR was used to isolate a full-length cDNA clone from λgt11 library. The fragment isolated from the yeast two hybrid system is indicated by an arrow and the three LIM domains are boxed.

	Endocytic codes	β-Galactosidase Activity LIM3/Enigma		Enigma	
		Glu	Gal	Glu	Gal
Exon 16/IR	QPDGPL**GPLY**ASSN**PEY**LSASD	–	100%	–	100%
APLA mut/IR	QPDGPL**APLA**ASSN**PEY**LSASD	–	–	–	–
APEA mut/IR	QPDGPL**GPLY**ASS**APEA**LSASD	–	60%	–	60%
double mut/IR	QPDGPL**APLA**ASS**APEA**LSASD	–	–	–	–
Exon 16/IGF1-R	NNSRLGN**GVLY**ASV**NPEY**FSASD	–	–	–	–
993-1022/EGFR	LIP**QQGFF**SSPSTSRTPLLSSLSATSNNS	–	–	–	–
730-780/LDLR	----------**NPVY**----------	–	–	–	–
1-61/TfR	----------**PLSYTRSF**----------	–	–	–	–
HTH/NNAYF	KQKVVKL**NNAYF**MFSHRI**NNAYF**KQKVVKL	–	3%	–	30%

Fig. 2 **Specific interaction of Enigma with endocytic codes of InsR.**
Interactions of various endocytic codes with LIM3 and full length Enigma were measured using the two hybrid system. Numbers refer to amino acid residues of the receptors with the defined endocytic codes indicated in bold. Yeast colonies were grown on X-gal indicator plates with glucose or with galactose. β-galactosidase activity was visualized and measured in solution. Activity quantitated by solution assay is expressed as percent relative to InsR exon 16.

A Tyrosine Random peptide Library

-4	-3	-2	-1	0	+1	+2	+3	+4
E(1.7)	E(1.7)	G(1.5)	P(2.0)	Y	G(1.3)	P(1.9)	V(2.4)	F(1.4)
I(1.4)	I(1.4)				H(1.3)		I(2.3)	H(1.4)
V(1.4)								

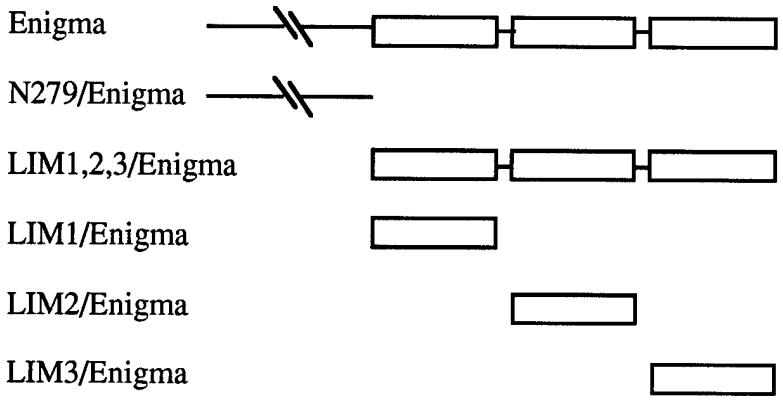
B Proline Random peptide Library

-4	-3	-2	-1	0	+1	+2	+3	+4	+5
X	X	G(2.1)	P	M(1.6)	G(1.4)	P	I(1.9)	Y(1.7)	A(1.7)
				F(1.6)			M(1.5)	F(1.4)	
				Y(1.6)			V(1.3)		
				I(1.4)					

C

Consensus:	G	P	Hyd	G	P	Hyd	Y/F	A
Exon16/InsR:	G	P	L	G	P	L	Y	A

Fig. 3 **Substrate specificity of LIM3 of Enigma detected by the degenerate peptide libraries.** Values in parentheses indicate the relative selectivity for the amino acids. X indicates no selectivity. The one-letter amino acid code is used. LIM3 of Enigma was expressed as a GST-fusion protein and purified by glutathione agarose. A tyrosine random peptide library or a proline random peptide library was presented to the purified GST-LIM3 of Enigma. The unbound peptides were washed away and the remaining peptide mixtures were sequenced and data was plotted.



	β-Galactosidase Activity			
	Ret/ptc2		Exon16/InsR	
	Glu	Gal	Glu	Gal
Enigma	6.9	945	8.0	894
N279/Enigma	4.0	8	6.7	8
LIM1,2,3/Enigma	6.5	664	9.5	776
LIM1/Enigma	5.4	22	7.5	14
LIM2/Enigma	8.1	227	4.3	13
LIM3/Enigma	3.4	9	11.7	420

Fig. 4 **Differential recognition of Ret/ptc2 and InsR by specific LIM domains of Enigma.**

β-galactosidase assays were performed on yeast expressing the LexA DNA binding domain Ret/ptc2 fusion protein and indicated B42 activation domain fusion proteins. β-galactosidase activity of each transformant was measured in solution. Numbers indicate the β-galactosidase activity (units/min/ml) quantitated by solution assays.

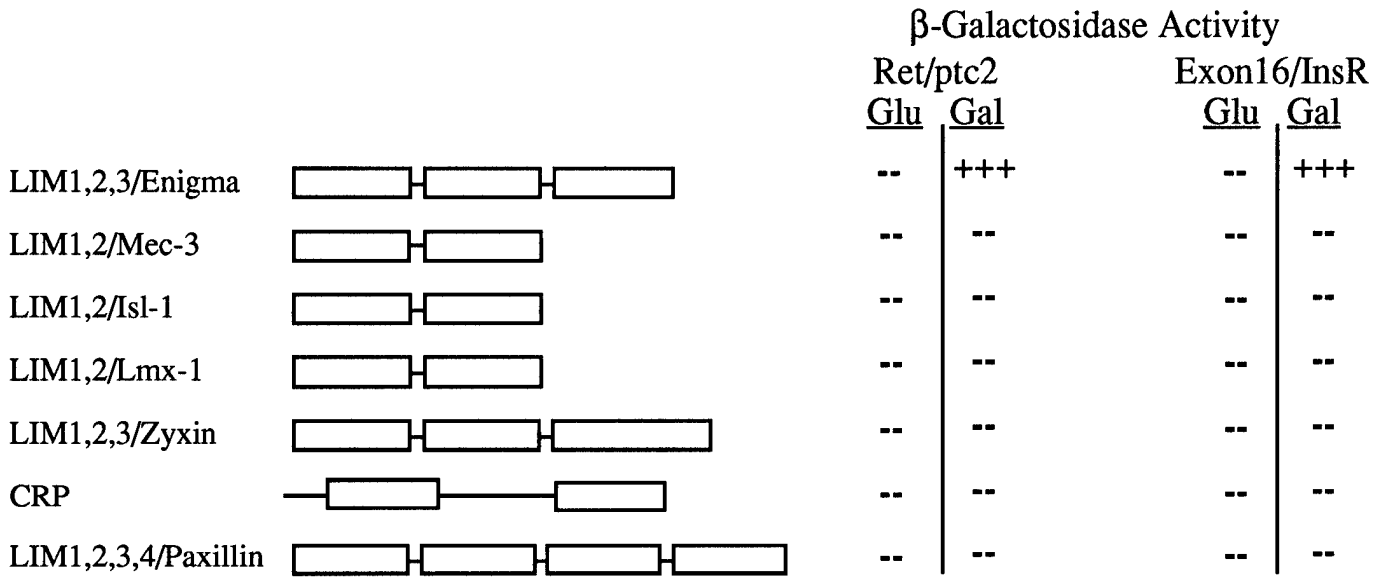


Fig. 5 **Comparison of Ret/ptc2 and exon 16 of InsR interaction with LIM domains.**

β-galactosidase activity of each transformant was visualized on X-gal plates. "+" indicates dark blue and "--" indicates white colonies.

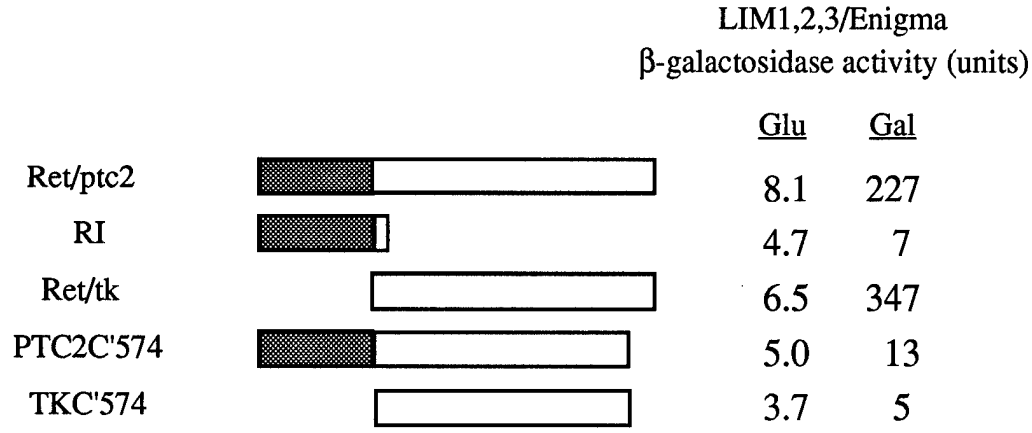


Fig. 6 **Identification of the carboxyl terminus of Ret/ptc2 required for interaction with the LIM domains of Enigma.**
Schematic representations of Ret/ptc2 and relative positions of deletion mutants. Interaction of the indicated construct with the LIM domains of Enigma was determined by β-galactosidase activity in the two hybrid system. Activity quantitated from solution assay is expressed as units/min/ml.

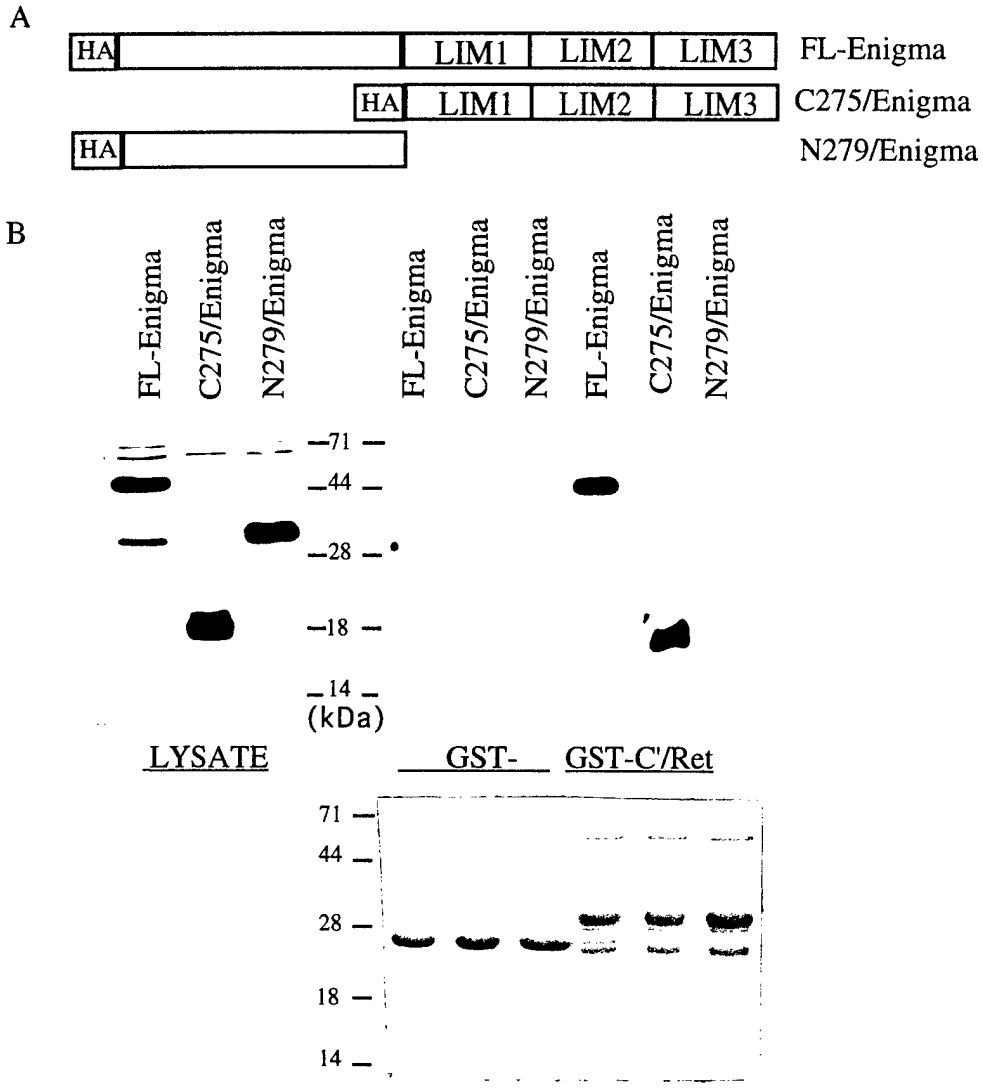


Fig. 7 **Direct association of Enigma with the carboxyl terminus of Ret.**

A, Schematic representation of Enigma fragments used in the binding assay. Full length and regions of Enigma were expressed as HA epitope-tag fusion proteins.

B, Binding of Enigma to the carboxyl terminal 61 amino acids (536-596) of Ret. 293 cells were transfected with indicated expression vectors, cell lysates were prepared 48 hours later and incubated with GST or GST-C' terminal Ret. Bound material was analyzed by western blotting with antibody to the HA epitope (upper). The bottom panel is a Coommassie blue-stained gel to quantitate GST and GST-C'/Ret used in the assays. The left panel shows the amount of protein in the lysate prior to interaction with GST-C'/Ret.

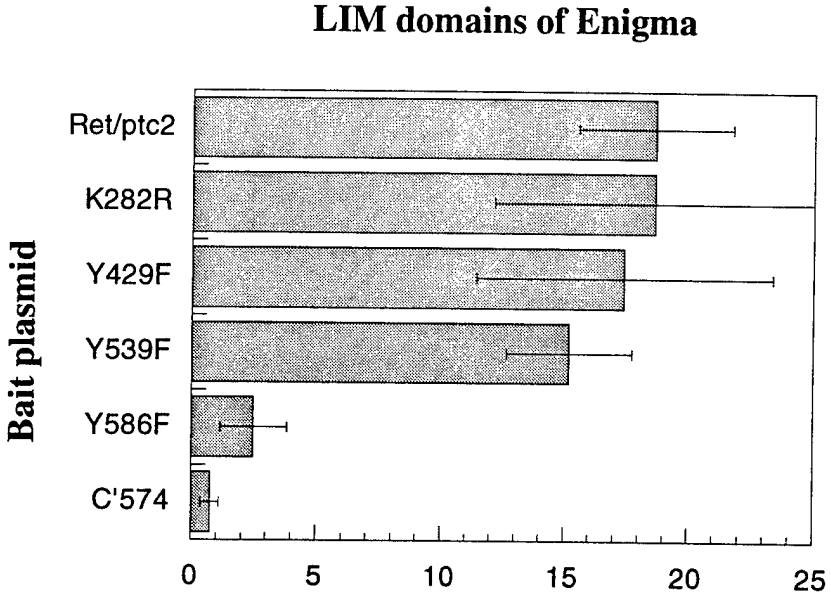


Fig. 8 **Tyr586 of Ret/ptc2 is required for Enigma-Ret interaction.**

Wild type and mutations of Ret/ptc2 were expressed as LexA fusion proteins and were tested for their ability to interact with the LIM domains of Enigma. K282R was a kinase-inactive mutant. Tyr429F, Tyr 539F and Tyr586F were point mutations of Tyr changed to Phe. C'574 was the carboxyl terminal 23 amino acids truncation of Ret/ptc2. 131 residues of mouse Enigma with 95% identity to the amino terminus of human Enigma was used. This Enigma fragment contained the LIM2 and LIM3. The β-galactosidase activity of yeast two-hybrid transformants was measured by solution assays.

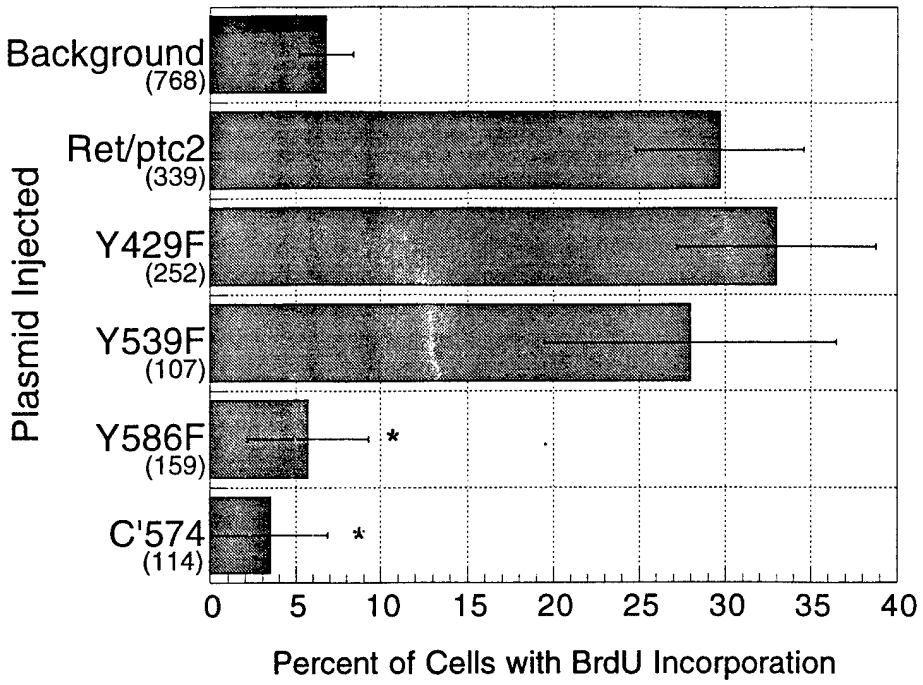


Fig. 9 **Mitogenic activity of Ret/ptc2 and mutants.**

Plasmids expressing either wild type Ret/ptc2 and various mutants were micoinjected into serum-starved fibroblasts and their ability to stimulate DNA synthesis were assessed by Immunofluorescence detection of 8-Bromodeoxyuridine (BrdU) incorporation. The numbers in parentheses are the total number of injected cells. Plasmids were injected at a concentration of 100μg/ml. (*) denotes a highly significant difference between cells injected with Ret/ptc2 or Y586F and Ret/ptc2 or C'574.

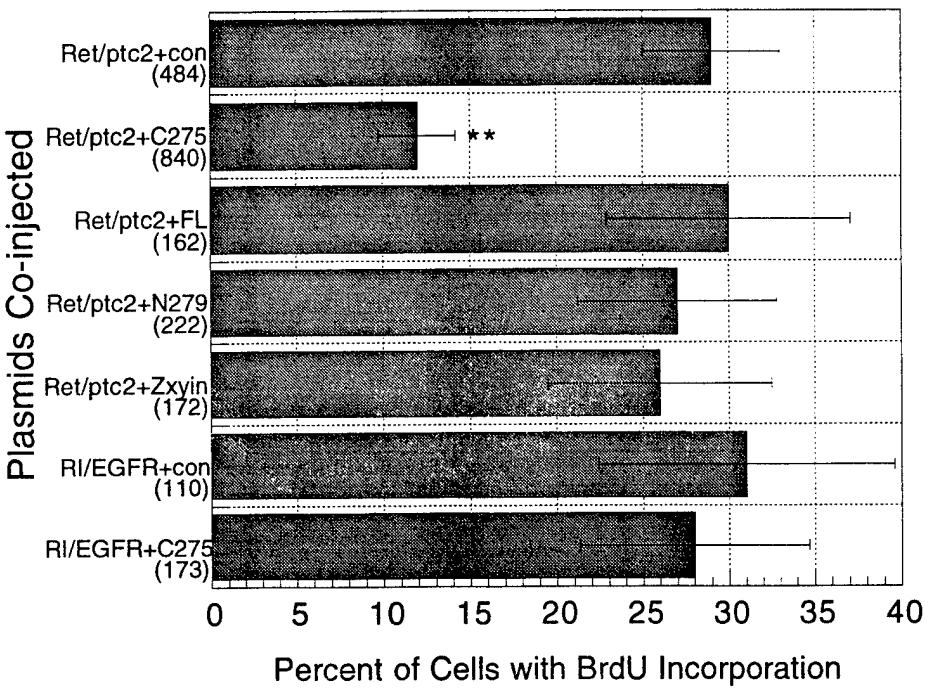


Fig. 10 **Inhibition of the mitogenic activity of Ret/ptc2 by co-expression with the LIM domains of Enigma.**

Serum-starved fibroblasts were microinjected with a mixture of two of the following expression plasmids: Ret/ptc2; RI/EGFR - analogous construct to Ret/ptc2 with EGFR intracellular domain in place of the Ret kinase; Con - control empty vector; C275 - the carboxyl terminal 275 amino acids of Enigma containing three LIM domains; FL - full length Enigma; N279 - the amino terminal 279 residues of Enigma lacking LIM domains; Zyxin - LIM domains from Zyxin, residues 339 to 542. In each case, Ret/ptc2 and RI/EGFR constructs were injected at 100µg/ml, while the othert constructs were presented at 200µg/ml. (**) denotes a highly significant difference between cells injected with Ret/ptc2+con or ret/ptc2+C275.

Bibliography

Name: Rui-Yun Wu

Address: Department of Growth and Development
University of California, San Francisco
Box 0640, C603
San Francisco, CA 94143
Tel: (415) 476-6081

Education:

B.S., Biology, July 1990
University of Science and Technology of China

Ph.D., Biology, March 1996
University of California at San Diego

Invention Disclosure:

[1] G. Gill and **R. Wu**, "Human Insulin Receptor Code Binding Protein", U.S. Patent, Serial Number 08/166,316, filed 12/13/93

Publications:

[1] **R. Wu** and G. Gill, "LIM Domain Recognition of a Tyrosine-containing Tight Turn", (1994) *J. Biol. Chem.*, **269**, 25085-25890

[2] K. Durick, **R. Wu**, G. N. Gill and S. S. Taylor, "Mitogenic Signaling by RET/ptc2 Requires Association with Enigma via a LIM Domains", (1996) *J. Biol. Chem.*, **271**, 12691-12694

[3] **R. Wu**, K. Durick, Z. Songyang, L. C. Cantley, S. S. Taylor and G. N. Gill, "Specificity of LIM Domain Interactions with Receptor Tyrosine Kinases", (1996) *J. Biol. Chem.*, **271**, 15934-15942

[4] C. Lazar, **R. Wu** and G. Gill, "Separation of Basolateral Targeting and Endosomal Sorting Sequences within the EGF Receptor", Submitted.

Published Abstracts:

[1] **R. Wu** and G. Gill, "A Novel Protein that Interacts Specifically with the Endocytic Codes of Human Insulin Receptor", American Society for Cell Biology Annual Meeting, New Orleans, Dec. 1993

[2] **R. Wu** and G. Gill, "LIM Domain Recognition of Tyrosine-containing Tight Turn", American Society for Cell Biology Annual Meeting, San Francisco, Dec. 1994

[3] C. Lazar, **R. Wu** and G. Gill, "Basolateral Sorting of Epidermal Growth Factor Receptors in MDCK Cells Requires Sequences Located in the Carboxyl Terminus", American Society for Cell Biology Annual Meeting, San Francisco, Dec. 1994

[4] **R. Wu** and G. Gill, "Differential Recognition of Endocytic Regions of the Insulin and Insulin-Like Growth Factor-1 Receptors by Specific LIM Domains", American Society for Clinical Investigation Annual Meeting, San Diego, May 1995

THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 269, No. 40, Issue of October 7, pp. 25085–25090, 1994
Printed in U.S.A.

LIM Domain Recognition of a Tyrosine-containing Tight Turn*

(Received for publication, July 14, 1994)

Rui-Yun Wu‡ and Gordon N. Gill§¶

From the Departments of ‡Biology and §Medicine, University of California San Diego, La Jolla, California 92093-0650

Endocytosis of cell surface receptors requires sequence "codes" consisting of tight turn structures with an essential Tyr or Phe residue. To determine mechanisms through which cells recognize this information, we utilized exon 16 of the human insulin receptor in the two-hybrid system to isolate a novel 455-amino acid cytoplasmic protein that contains two LIM domains within its carboxyl terminus. Mutational analyses indicate that one of the Cys-rich Zn^{2+} binding LIM domains specifically recognizes active but not inactive endocytic codes contained in exon 16. These findings suggest that LIM domain structures in proteins provide molecular recognition of Tyr-containing tight turn structures.

Ligand-induced internalization of insulin receptors (InsR) requires intrinsic protein tyrosine kinase activity and sequences encoded by exon 16 (1). Exon 16, which encodes 22 amino acids of the submembrane domain of the β subunit of InsR, contains two endocytic sequences: a strong code, ^{950}GPLY, and a weak code, ^{957}NPEY (2, 3). Both codes form Tyr-containing tight turns in solution (4). Mutation of GPLY to APLA and NPEY to APEA reduced InsR internalization to 32 and 87%, respectively, of the rate of wild-type InsR and mutation of both codes abolished ligand-induced internalization (2); the Y → A change disrupted the tight turn structure (4). A Tyr-containing tight turn structure is the essential recognition motif present in endocytic codes of many receptors (5–8). The requirement for kinase activity in signaling receptors reflects in large part conformational changes necessary to expose these codes (9, 10).

Because the mechanisms by which cells recognize endocytic codes is unknown, we used the two-hybrid system (11) to isolate a protein that specifically interacts with exon 16 of human InsR. This novel 455-amino acid protein contains two LIM domains at its COOH terminus. LIM (*lin*-11 (12), *isl*-1 (13), *mec*-3 (14)) was initially identified as a Cys-rich domain located at the NH_2 terminus of homeodomain proteins involved in development. LIM domains have also been identified in zyxin and cCRP that are associated with focal adhesions and the cytoskeleton (15, 16). LIM domains, which bind Zn^{2+} (17, 18) and are often present in more than one copy within proteins, are proposed to function in protein-protein interactions both in homeodomain proteins that act in the nucleus and in proteins that act in the cytoplasm and with the cytoskeleton (12, 15).

The protein which, based on its code recognition properties, we term Enigma (19), interacted with InsR exon 16 through the COOH-terminal LIM domain (LIM2). Neither LIM1 of Enigma nor the LIM domains of *mec-3* interacted with exon 16. There was specificity in target recognition as well because mutation of GPLY to APLA in InsR exon 16 abolished interaction, and Enigma failed to interact with fragments of the low density lipoprotein receptor (LDLR)[1] (20), transferrin receptor (TfR) (21), and epidermal growth factor receptor (EGFR) (9) that contain endocytic codes. To investigate features of LIM domain interaction with proteins, a "synthetic" sequence containing NNAYF, a motif that was shown to function in endocytosis of mutant EGFR (9), was tested for interaction using the two-hybrid system. The NNAYF code was recognized by both LIM domains of Enigma and by LIM2 of *mec-3*, suggesting that LIM domains recognize proteins through interaction with exposed tyrosines located in tight turns and possibly other structures.

EXPERIMENTAL PROCEDURES

The Two-hybrid System—Reagents for the two-hybrid system were obtained from Dr. Roger Brent (22, 23). PCR was used to isolate exon 16 from InsR cDNA (24), and this was used to create a LexA-exon 16 fusion that was transcriptionally inactive but able to bind LexA operators (22, 23). This "bait" was used to screen a HeLa cell cDNA fusion library in the vector pJG4–5, which was introduced into the yeast strain EGY 48(MATα, His3, Ura3–52, Trp1, Leu2:pLEU2LexAop6). JG4–5 contains a galactose-inducible promoter, SV40 T antigen nuclear localization sequence, B42 activation domain, and the HA1 epitope. 10^6 transformants were induced with galactose (20-fold expansion) and 398 Leu$^+$ colonies isolated. Three colonies containing identical inserts were confirmed by galactose-inducible LexA-operator-driven β-galactosidase activity. Because the oligo(dT)-prepared HeLa cell library would yield one-third of the cDNAs in the correct reading frame, Enigma is represented in ~1 in 3×10^5 cDNAs of the library.

PCR primers were used to isolate DNA fragments encoding the endocytic codes of LDLR, TfR, EGFR, and IGF-1R. cDNAs containing mutant InsR exon 16 sequences were obtained from Dr. Donald McClain (2) and mutant exon 16 sequences were isolated by PCR. The NNAYF code was made by annealing two complimentary pairs of oligonucleotides and ligating them together into pEG202. KQKVVKL and MFSHRI correspond to helices B and D, respectively, from the crystallographic structure of protein kinase A (25). DNAs were fused in frame to the DNA binding domain of the LexA protein of the yeast vector pEG202 (22, 23).

LIM2 of Enigma was isolated from the original library screening based on interaction with exon 16 of InsR. Holo-Enigma was made from the full-length cDNA isolated from the SK-N-MC cell λgt11 library and cloned in-frame to the yeast library plasmid pJG4–5. cDNA clones corresponding to various LIM domains of Enigma and *mec*-3 were isolated by PCR and fused in-frame to the pJG4–5 vector. Sequences of PCR primers and of oligonucleotides are available from the authors. Plasmids that direct the synthesis of LIM domains and holo-Enigma were introduced into EGY48/1840 that contained different LexA fusion endocytic code proteins using the lithium acetate procedure (26). All constructs were confirmed by dideoxynucleotide sequencing (U. S. Biochemical Corp.), and the expression of fusion proteins of the appropriate size was confirmed by Western blotting using α-HA1 antibody (BAbCO).

Transformants were grown on selective medium (Ura$^-$ His$^-$ Trp$^-$). Cells from individual transformants were streaked onto Ura$^-$ His$^-$ Trp$^-$

* This study was supported by National Institutes of Health Grant PO1 CA58689 and American Cancer Society Grant CB-3P. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EMBL Data Bank with accession number(s) L35240.

¶ To whom correspondence should be addressed: UCSD School of Medicine, Division of Endocrinology/Metabolism 0650, La Jolla, CA 92093-0650. Tel.: 619-534-4310; Fax: 619-534-8193.

[1] The abbreviations used are: LDLR, low density lipoprotein receptor; PCR, polymerase chain reaction; TfR, transferrin receptor; EGFR, epidermal growth factor receptor; IGF-1R, insulin-like growth factor-1 receptor; GST, glutathione *S*-transferase.

a

```
GAATTCCCGTTGCTGTCGCCCAACGAGGCTCCCTGGAGCCGACGCAGAGCAGCGCCCTGGCCGGGCCAAGCAGGAGCCGGCATC 84
ATGGATTCCTTCAAAGTAGTGCTGGAGGGGCCAGCACCTTGGGGGTTCCGGCTGCAAGGGGGCAAGGACTTCAATGTGCCCCTCTCCATT 174
 M  D  S  F  K  V  V  L  E  G  P  A  P  W  G  F  R  L  Q  G  G  K  D  F  N  V  P  L  S  I  30
TCCCGGCTCACTCCTGGGGGCAAAGCGGCGCAGGCCGGAGTGGCCGTGGGTGACTGGGTGCTGAGCATCGATGGCGAGAATGCGGGTAGC 264
 S  R  L  T  P  G  G  K  A  A  Q  A  G  V  A  V  G  D  W  V  L  S  I  D  G  E  N  A  G  S  60
CTCACACACATCGAAGCTCAGAACAAGATCCGGGCCTGCGGGGAGCGCCTCAGCCTGGGCCTCAGCAGGGCCCAGCCGGTTCAGAGCAAA 354
 L  T  H  I  E  A  Q  N  K  I  R  A  C  G  E  R  L  S  L  G  L  S  R  A  Q  P  V  Q  S  K  90
CCGCAGAAGGCCTCCGCCCCCGCCGCGGACCCTCCGCGTACACCTTTGCACCCAGCGTCTCCCTCAACAAGACGGCCCAACCCTTTGGGC 444
 P  Q  K  A  S  A  P  A  A  D  P  P  R  T  P  L  E  P  A  S  P  S  T  R  R  P  N  P  L  G  120
CCCCCCGCCCCTGACAGCCCCCCGCAGCAGAATGGACAGCCGCTCCGACCGCTGGTCCCAGATGCCAGCAAGCAGCGGCTGATGGAGAAC 534
 P  P  A  P  D  S  P  P  Q  Q  M  G  Q  P  L  R  P  L  V  P  D  A  S  K  Q  R  L  M  E  N  150
ACAGAGGACTGGCGGCCGCGGCCGGGACAGGCCAGTCGCGTTCCTTCCGCATCCTTGCCCACCTCACAGGCACCGAGTTCATGCAAGACC 624
 T  E  D  W  R  P  R  P  G  Q  A  S  R  V  P  S  A  S  L  P  T  S  Q  A  P  S  S  C  K  T  180
CCGGATGAGGAGCACCTGAAGAAATCAAGCCAGGTGCCAGACAGAAGCCCCAGCCCCAGCCTCATCTACACCCCAGGAGCCCTGGCCTGG 714
 P  D  E  E  H  L  K  K  S  S  Q  V  P  D  R  S  P  S  P  S  L  I  Y  T  P  G  A  L  A  W  210
CCCTACCGCCCCCAGCCCTACCAGCCGCCCGCCCTGGGCTGTGGACCCTGCGTTTGCCGAGCGCTATGCCCCGGACAAAACGAGCACAGT 804
 P  Y  R  P  Q  P  Y  Q  P  P  A  L  G  C  G  P  C  V  C  R  A  L  C  P  G  Q  N  E  H  S  240
GCTGACCCACACAGCCAGCCAGCCACGCCCACGCCGCTGCAGAGCCGCACCTCCATTGTGCAGGCAGCTGCCGGAGGGGTGCCAGGAGGG 894
 A  D  P  H  S  Q  P  A  T  P  T  P  L  Q  S  R  T  S  I  V  Q  A  A  A  G  G  V  P  G  Q  270
GGCAGCAACAACGGCAAGACTCCCGTGTGTCACCAGTGCCACAAGGTCATCCGGGGCCGCTACCTGGTGGCGCTGGGCCACGCGTACCAC 984
 G  S  N  N  G  K  T  P  V  C  H  Q  C  H  K  V  I  R  G  R  Y  L  V  A  L  G  H  A  Y  H  300
CCGGAGGAGTTTGTGTGTAGCCAGTGTGGGAAGGTCCTGGAAGAGGGTGGCTTCTTTGAGGAGAAGGGCGCCATCTTCTGCCCACCATGC 1074
 P  E  E  F  V  C  S  Q  C  G  K  V  L  E  E  G  G  F  F  E  E  K  G  A  I  F  C  P  P  C  330
TATGACGTGCGCTATGCACCCAGCTGTGCCAAGTGCAAGAAGAAGATTACAGGCGAGATCATGCACGCCCTGAAGATGACCTGGCACGTG 1164
 Y  D  V  R  Y  A  P  S  C  A  K  C  K  K  K  I  T  G  E  I  M  H  A  L  K  M  T  W  H  V  360
CACTGCTTTACCTGTGCTGCCTGCAAGACGCCCATCCGGAACAGGGCCTTCTACATGGAGGAGGGCGTGCCCTATTGCGAGCGAGACTAT 1254
 H  C  F  T  C  A  A  C  K  T  P  I  R  N  R  A  F  Y  M  E  E  G  V  P  Y  C  E  R  D  Y  390
GAGAAGATGTTTGGCACGAAATGCCATGGCTGTGACTTCAAGATCGACGCTGGGGACCGCTTCCTGGAGGCCCTGGGCTTCAGCTGGCAT 1344
 E  K  M  F  G  T  K  C  H  G  C  D  F  K  I  D  A  G  D  R  F  L  E  A  L  G  F  S  W  H  420
GACACCTGCTTCGTCTGTGCGATATGTCAGATCAACCTGGAAGGAAAGACCTTCTACTCCAAGAAGGACAGGCCTCTCTGCAAGAGCCAT 1434
 D  T  C  F  V  C  A  I  C  Q  I  N  L  E  G  K  T  F  Y  S  K  K  D  R  P  L  C  K  S  H  450
GCCTTCTCTCATGTGTGAGCCCCTTCTGCCCACAGCTGCCGCGGTGGCCCCTAGCCTGAGGGGCCTGGAGTCGTGGCCCTGCATTTCTGG 1524
 A  F  S  H  V 455
GTAGGGCTGGCAATGGTTGCCTTAACCCTGGCTCCTGGCCCGAGCCTGGGCTCCCTGGCCCTGCCCCACCCACCTTATCCTCCCACCCCA 1614
CTCCCTCCACCACCACAGCACACCGGTGCTGGCCACACCAGCCCCCTTTCACCTCCAGTGCCACAATAAACCTGTACCCAGCTGAAAAAA 1704
AAAAAAAAAAAAAAAACTCGAG 1725
```

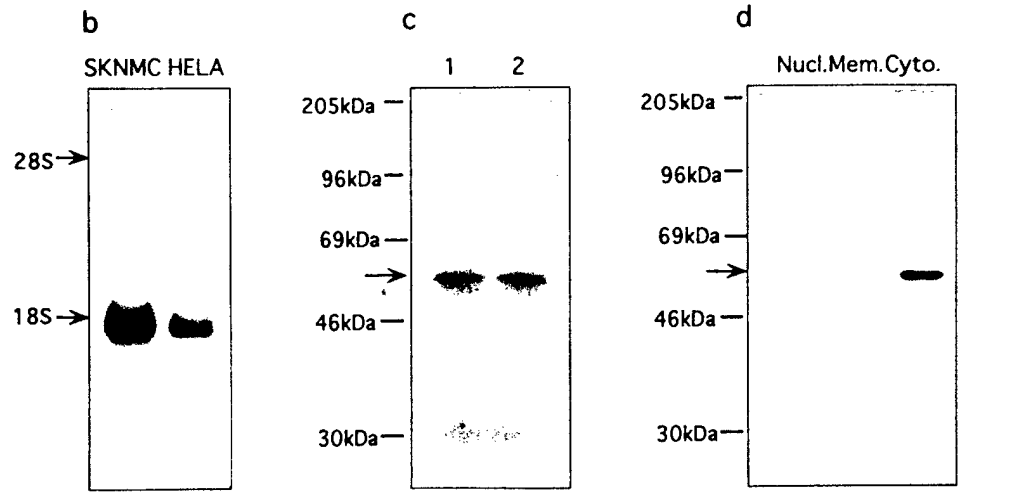
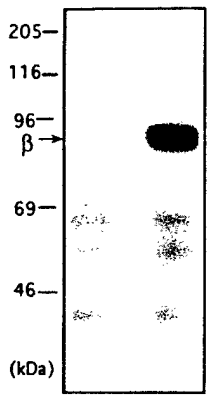


FIG. 1. **Identification and sequence of Enigma.** *a*, the nucleotide and amino acid sequence of the 455-residue protein. The COOH terminus of Enigma isolated from a HeLa cell library using the two-hybrid system to detect interaction with the 22 amino acids encoded by exon 16 of InsR was used to isolate a full-length cDNA clone from a λgt11 cDNA library prepared from SK-N-MC cells. The COOH-terminal fragment isolated by the two-hybrid screen is marked by an *arrow*; the LIM domains, LIM1 and LIM2, are *boxed*. A Gly-rich region with homology to a nucleotide binding site is *underlined*. *b*, Northern blot analysis of 2 μg of poly(A)$^+$ RNA isolated from HeLa and SK-N-MC cells. The 7-kDa COOH terminus of Enigma excised from pJG4–5 was used as a hybridization probe. *c*, two full-length cDNA clones were placed in pBS in both orientations. mRNA was transcribed using T7 and T3 polymerase and translated *in vitro* in rabbit reticulocyte lysates. The ^{35}S-labeled reaction products were analyzed on SDS-polyacrylamide gel electrophoresis gels and autoradiographed. *d*, SK-N-MC cells were fractionated by standard procedures and aliquots analyzed by Western blotting with α Enigma. Antibody was prepared by immunization of rabbits with the COOH-terminal 7-kDa fragment of Enigma made as a GST fusion protein in *E. coli*. The IgG fraction was enriched by ammonium sulfate fractionation and immunopurified on Sepharose-immobilized 7-kDa fragment produced in *E. coli* as a poly(His) fusion protein. Detection was with HRP-conjugated donkey α rabbit Ig (Amersham Corp.).

5-bromo-4-chloro-3-indoyl β-D-galactoside plates containing glucose or galactose for determination of β-galactosidase activity. Quantitative analysis of β-galactosidase activity was performed in liquid culture. Background levels of β-galactosidase activity were calculated using a buffer only blank or with LexA fused to *Drosophila bicoid* which is a standard negative in the two-hybrid system (22, 23). Activity was cal-

a



b

	Endocytic Codes	Activity: Lim2 Glu	Lim2 Gal	Holo Enigma Glu	Holo Enigma Gal
Exon 16/IR	QPDGPL**GPLY**ASS**NPEY**LSASD		100%		100%
APLA mut/IR	QPDGPL**APLA**ASS**NPEY**LSASD		-		-
APEA mut/IR	QPDGPL**GPLY**ASS**APEA**LSASD		60%		80%
double mut/IR	QPDGPL**APLA**ASS**APEA**LSASD		-		-
Exon 16/IGF1-R	NNSRLGN**GVLY**ASV**NPEY**FSASD		-		-
993-1022/EGFR	LIP**QQGFF**SSPSTSRTPLLSSLSATSNNS		-		-
993-1022/NNAYF	LIP**NNAYF**SSPSTSRTPLLSSLSATSNNS		-		-
730-780/LDLR	----------**NPVY**----------		-		-
1-61/TfR	----------PLS**YTRF**---------		-		-
HTH/NNAYF	KQKVVKL**NNAYF**MFSHRI**NNAYF**KQKVVKL		3%		30%

FIG. 2. **Specificity of interaction of Enigma with endocytic codes of InsR.** *a*, *in vitro* binding of InsR to LIM2 of Enigma. Partially purified InsR was incubated with glutathione-agarose-immobilized GST or GST-7kDa Enigma, and the bound protein was detected by Western blotting using antibody specific for the β subunit of InsR. *b*, interaction of various endocytic codes with LIM2 and holo-Enigma measured using the two-hybrid system. *Numbers* refer to amino acid residues of the indicated receptors with the defined endocytic codes indicated in bold. Yeast colonies were grown on 5-bromo-4-chloro-3-indoyl β-D-galactoside indicator plates with glucose or with galactose. β-Galactosidase activity was visualized and measured in solution. Activities quantitated from solution assays are expressed as percent relative to InsR exon 16.

culated using triplicates in three separate assays according to: units = 1000 ($OD_{420}/OD_{600}tv$) when t = time and v = volume. Data are presented as percentages relative to interaction with exon 16 of InsR which was set to 100%.

Cloning and Sequencing of Enigma—The 500-base pair insert isolated in the two-hybrid screen using the LexA-exon 16 fusion was used as a probe to screen a λgt11 cDNA library prepared from SK-N-MC cells provided by Dr. Eric Turner, University of California, San Diego. Eight positive clones were isolated, four of which contained the full-length cDNA based on Northern blot analysis that identified a single 1.73-kilobase mRNA. The complete cDNA was sequenced using both transposon insertions (27) and standard primers in the dideoxynucleotide chain termination procedure. The complete sequence was analyzed for overall homology and for domain homology using the programs BLAST and QUEST.

Interaction of InsR with GST-LIM2—The 500-base pair fragment which encodes the 7-kDa COOH-terminal fragment of Enigma was excised from pJG4–5 by *Eco*RI and *Xho*I digestion and cloned into pBS (Stratagene) at these sites. The 500-base pair fragment was excised from pBS with *Bam*HI and *Xho*I and ligated in frame into pGEX-KG (Pharmacia). This plasmid was used to produce GST-7kDa in bacteria and the fusion protein isolated on a glutathione-agarose column. InsR was purified from HIRc cells by wheat germ lectin affinity chromatography and incubated with either immobilized GST or GST-7kDa. Bound protein was eluted, separated by SDS-polyacrylamide gel electrophoresis, and β subunits of InsR detected by Western blotting with α InsR antibody.

RESULTS

Using the two-hybrid reagents developed by Roger Brent and colleagues (22, 23) a 500-base pair HeLa cell cDNA was identified that encoded a 7-kDa protein fragment that interacted with the 22 amino acids of exon 16. This cDNA fragment was used to isolate a full-length cDNA from a λgt11 cDNA library prepared from SK-N-MC cells (Fig. 1). The cDNA contains an open reading frame encoding a 455-amino acid protein. The cDNA corresponds to a single mRNA species of 1.7 kilobase determined by Northern blotting (Fig. 1*b*), contains a Kozak sequence at the initiator Met, and the reading frame confirms that determined for the COOH-terminal 7-kDa fragment produced as a fusion protein by plasmid pJG4–5 in the two-hybrid screen. *In vitro* translation yielded a 55-kDa protein as expected from the open reading frame (Fig. 1*c*). Moreover, when full-length Enigma was placed into pJG4–5 it specifically interacted with exon 16 in the two-hybrid system (see Fig. 2). Western blotting with an affinity-purified antibody generated against the 7-kDa fragment confirmed a single cytoplasmic ~55-kDa protein in SK-N-MC (and HeLa) cells in agreement with the open reading frame and *in vitro* translation (Fig. 1*d*).

Although Enigma is a novel protein not present in the data banks, the 7-kDa COOH terminus that interacted with exon 16 in the two-hybrid system consists predominantly of a LIM domain. Enigma contains two LIM domains in its COOH terminus which are highly related (38% amino acid identity, see Fig. 3).

Holo-InsR purified from HIRc cells by wheat germ lectin affinity chromatography was bound to the 7-kDa COOH terminus of Enigma expressed as a GST fusion protein but not to GST alone (Fig. 2*a*). To determine whether Enigma recognized the active endocytic codes of InsR, the two-hybrid system was used to measure interaction of Enigma with mutant exon 16 sequences. Both the 7-kDa COOH-terminal fragment containing LIM2 and holo-Enigma were used. A 2-amino acid change in the strong endocytic code (GPLY to APLA) almost completely abolished interaction (Fig. 2*b*). A 2-amino acid change in the weaker endocytic code (NPEY to APEA) decreased interaction but did not abolish it. No interaction occurred when both mutations were present. There is thus strong concordance between effects of these point mutants in InsR *in vivo* (2–4) and *in vitro* interactions between Enigma and exon 16.

Although many receptors internalize via coated pits, specificity is implied by observations that EGFR do not compete with TfR for internalization (28) nor do InsR compete with IGF-II/mannose 6-phosphate receptors for internalization (29). To investigate specificity, we tested the interaction of LIM2 and holo-Enigma with endocytic codes from other receptors. As shown in Fig. 2*b*, no interaction was detected using the two-hybrid system with the cytoplasmic domains that contain the endocytic codes of IGF1-R (30), LDL receptor (LDLR) (20), TfR

a

```
Enigma:1  CAKCKKKI.TGE..I.MHA.LKMTWHVHCFTCAACKTPI...RNRAFYMEEGVPYCERD
Enigma:2  CHGCDFKIDAGDRFLE..A.LGFSWHDTCFVCAICQ..INL.EGKTFYSKKDRPLCKSH
Mec-3:1   CNCCNEQI..YDRYIYRMDNR..SYHENCVKCTICESP..LAE.KCFW.KNGRIYCSQH
Mec-3:2   CAGCKKGVSPTDMV.YKLKAGL.VFHVECHCCSLCGRHLSPGEQILVDDTMMTVSCMSH

          CX2C.....X16-19.....HX2CX2CX2C.....X16-20.....CX2H,D
```

b

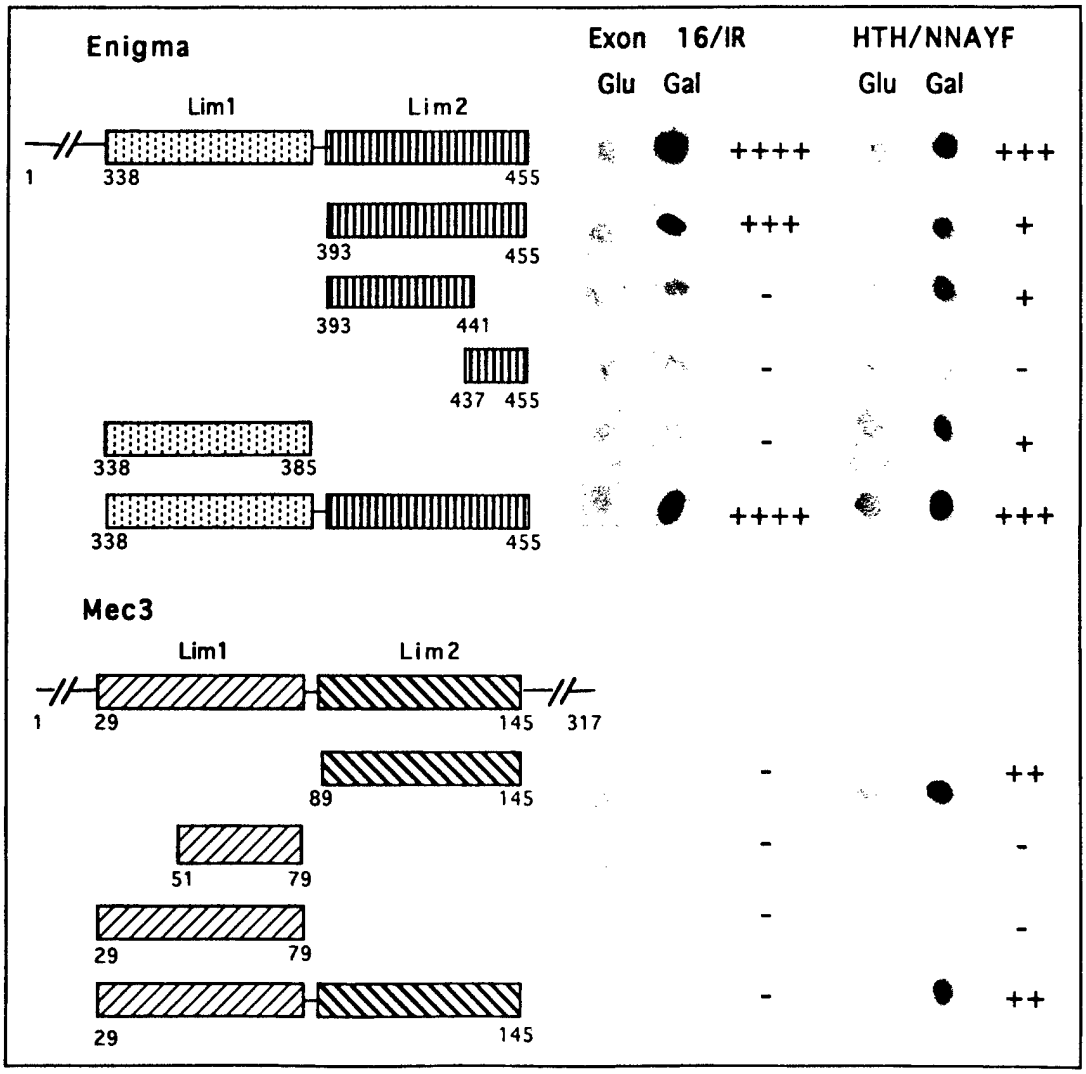


FIG. 3. **The ability of LIM domains of Enigma and *mec-3* to interact with exon 16 of InsR and the NNAYF synthetic code.** *a*, amino acid sequence alignment of the LIM domains of Enigma and *mec-3*. *b*, interaction of the indicated LIM domains with exon 16 of InsR and with the NNAYF synthetic code as measured by β-galactosidase activity in the two-hybrid system. *Numbers* indicate amino acid residues of the LIM domains used.

(21), or EGFR (9). Interaction between LIM2 of Enigma and the GPLY strong endocytic code of InsR thus showed high specificity. An interaction was detected with a sequence containing a synthetic "idealized" endocytic code (NNAYF) that was demonstrated to function in EGFR (9). Although NNAYF was not recognized when substituted for a single endocytic code of EGFR, a significant interaction was observed when it was expressed in two copies in a predicted helix-turn-helix as occurs in native EGFR where two endocytic codes are separated by a predicted helical region (9, 31). The interaction of holo-Enigma was stronger than the interaction of LIM2 with the NNAYF code.

Specificity for LIM2 recognition of exon 16 was examined using LIM1 of Enigma and LIM1 and LIM2 of *mec-3* (14). Fig. 3*a* shows the alignment of the LIM domains of Enigma and *mec-3*, and Fig. 3*b* shows that neither LIM1 of Enigma nor LIM domains 1 or 2 of *mec-3* recognized exon 16. The specificity of Enigma for exon 16 of InsR thus resides in LIM2. Similar to the results seen with holo-Enigma, LIM1 and the LIM domains of *mec-3* failed to recognize endocytic codes of IGF1-R, LDLR, TfR, and EGFR (data not shown).

The ability of LIM2 of Enigma to specifically recognize the active endocytic code of exon 16 of InsR provided an initial approach to determine the sequence necessary to constitute an

active and specific LIM domain. LIM2 was divided into a core and a COOH-terminal extension based on homology among LIM domains (12, 15) (Fig. 3*a*). As shown in Fig. 3*b*, neither part alone interacted with exon 16, indicating that only the complete LIM2 structure containing ~50 amino acids and two Zn^{2+} coordination sites (Cys^3-His and Cys^4) (16) constitutes a fully functional LIM domain.

The ability of Enigma to recognize NNAYF, a synthetic code in the context of surrounding helices, suggested this might serve as a general target for LIM domains. As shown in Fig. 3*b*, both LIM1 and LIM2 of Enigma and LIM2 of *mec-3* recognized this structure equally well. A portion of LIM2 of Enigma (residues 393–441), which failed to recognize exon 16, weakly recognized the NNAYF synthetic code but LIM1 of *mec-3* did not. This suggests that tight turns containing an aromatic residue are a more general motif for LIM domain binding when presented in an appropriate context. Variations in the sequence of tight turns and of LIM domains provide specificity for protein-protein interactions.

DISCUSSION

The ability to specifically recognize the active endocytic code of InsR via a LIM domain fulfills the first property of the endocytic mechanism predicted from study of the kinetics of ligand-induced saturable high affinity endocytosis of receptors, that of recognition (9), but additional criteria are necessary to demonstrate function in this process. Enigma recognized the endocytic codes of exon 16 of InsR proportional to the strength of the codes in holo-InsR *in vivo* (2–4), but failed to recognize other endocytic codes. Specificity resided not only in the tyrosine-containing tight turn but in the LIM domain as shown by the inability of LIM1 of Enigma and both LIM1 and LIM2 of *mec-3* to interact with exon 16. Although these observations indicate specificity for both partners to the interaction, they do not prove a function for Enigma which may act in ways unrelated to insulin action. The two-hybrid system has high sensitivity for detecting protein interactions (11), and it remains possible that physiologically relevant interactions with exon 16 and with Enigma will be quite different from those described. The interactions between exon 16 of InsR and LIM2 of Enigma do nonetheless provide evidence for a novel protein interaction mechanism. NMR analysis of peptides corresponding to exon 16 of InsR indicates that the GPLY sequence is a tight turn structure (4). The NMR structure of the LIM domain of CRP indicates it is largely composed of β sheets surrounding a hydrophobic core with the a Zn^{2+} atom coordinated at each end (32). The present studies indicate that a LIM domain structure can interact with a tyrosine-containing tight turn. The finding that the "generic" code NNAYF was recognized equally well by both LIM domains of Enigma and by LIM2 of *mec-3* suggests that LIM domains provide the structural basis for recognition of other tyrosine-containing tight turn structures.

NNAYF was originally chosen to mimic or change the native EGFR sequence QQGFF (9). In the context of mutant EGFR it functions as an endocytic code (9) and is the only sequence found that restored endocytosis to kinase-inactive EGFR.[2] Although the structure of NNAYF is not known, its placement between sequences corresponding to helices in protein kinase A is expected to provide exposure perhaps as a helix-turn-helix. The interaction of several LIM domains with this structure provides support for the hypothesis that LIM domains recognize exposed tyrosine-containing motifs. LIM domains may additionally recognize DNA or other structures (32).

LIM domains are proposed to function in protein-protein recognition in a variety of contexts including gene transcription and development (12–14, 33–35) and in cytoskeletal interaction (15, 16). Like SH2 and SH3 domains (36), LIM domains are distributed in a variety of proteins and are proposed to function in assembly and disassembly of protein complexes. Interactions between two LIM domain-containing proteins, zyxin and cCRP, suggest that self-association may prove important (15, 16) in addition to recognition of motifs such as tight turns. LIM domains in the transcription factor *lmx*-1 are necessary for specific functional interaction with sh*Pan-1* which lacks LIM domains (37), supporting the idea that LIM domains interact with structural motifs distinct from other LIM domains. Although the spacing of Cys and His residues is the central structural motif of LIM domains providing for coordination of Zn^{2+} to create structures that bear homology to Zn^{2+} fingers, there is significant variation in the intervening amino acid sequence (32). SH2 and SH3 domains, although widely distributed in proteins, exhibit specificity in interactions (38), suggesting this must be true of LIM domains to account for their wide distribution in proteins of differing function. Multiple copies of each component, *i.e.* multiple tight-turn structures within one protein and multiple LIM domains in another protein, could increase the strength of interactions and broaden recognition properties.

Acknowledgments—We gratefully acknowledge generous help with the two-hybrid system from Roger Brent, Grace Gill, and Scott Emr. Eric Turner, Don McClain, Serge Lichsteiner, Steven Wiley, Lynn Sealey, Jerry Olefsky, Grahame Bell, and Donna Palm provided essential reagents.

[2] H. S. Wiley and G. N. Gill, unpublished data.

REFERENCES

1. Theis, R. S., Webster, N. J., and McClain, D. A. (1990) *J. Biol. Chem.* **265,** 10132–10137
2. Rajagopalan, M., Neidigh, J. L., and McClain, D. A. (1991) *J. Biol. Chem.* **266,** 23068–23073
3. Berhanu, P., Ibrahim-Schneck, R. H., Anderson, C., and Wood, W. M. (1991) *Mol. Endocrinol.* **5,** 1827–1835
4. Backer, J. M., Shoelson, S. E., Weiss, M. A., Jua, Q. X., Chatham, R. B., Harig, E., Cahill, D. C., and White, M. F. (1992) *J. Cell Biol.* **118,** 831–839
5. Collawn, J. F., Stangel, M., Kuhn, L. A., Esekogwu, V., Jing, S., Trowbridge, I. S., and Tainer, J. A. (1990) *Cell* **63,** 1061–1072
6. Bansal, A., and Gierasch, L. M. (1991) *Cell* **67,** 1195–1201
7. Eberle, W., Sander, C., Klaus, W., Schmidt, B., von Figura, K., and Peters, C. (1991) *Cell* **67,** 1203–1209
8. Trowbridge, I. S., Collawn, J. F., and Hopkins, C. R. (1993) *Annu. Rev. Cell Biol.* **9,** 129–161
9. Chang, C.-P., Lazar, C. S., Walsh, B. J., Komuro, M., Collawn, J. F., Kuhn, L. A., Tainer, J. A., Trowbridge, I. S., Farquhar, M. G., Rosenfeld, M. G., Wiley, H. S., and Gill, G. N. (1993) *J. Biol. Chem.* **268,** 19312–19320
10. Cadena, D. L., Chan, C., and Gill, G. N. (1994) *J. Biol. Chem.* **269,** 260–265
11. Fields, S., and Song, O.-K. (1989) *Nature* **340,** 245–246
12. Freyd, G., Kim, S. K., and Horvitz, H. R. (1990) *Nature* **344,** 876–879
13. Karlsson, O., Thor, S., Njorberg, T., Ohlsson, H., and Edlund, T. (1990) *Nature* **344,** 879–882
14. Way, J. C., and Chalfie, M. (1988) *Cell* **54,** 5–16
15. Sadler, I., Crawford, A. W., Michelsen, J. W., and Beckerle, M. C. (1992) *J. Cell Biol.* **119,** 1573–1587
16. Crawford, A. W., Pino, J. D., and Beckerle, M. C. (1994) *J. Cell Biol.* **124,** 117–127
17. Michelsen, J. W., Schmeichel, K. L., Beckerle, M. C., and Winge, D. R. (1993) *Proc. Natl. Acad. Sci. U. S. A.* **90,** 4404–4408
18. Michelsen, J. W., Sewell, A. K., Louis, H. A., Olsen, J. I., Davis, D. R., Winge, D. R., and Beckerle, M. C. (1994) *J. Biol. Chem.* **269,** 11108–11113
19. Welchman, G. (1982) *The Hut Six Story-Breaking the Enigma Codes*, McGraw Hill Book Co., New York
20. Chen, W. J., Goldstein, J. L., and Brown, M. S. (1990) *J. Biol. Chem.* **265,** 3116–3123
21. Jing, S., Spencer, T., Miller, K., Hopkins, C., and Trowbridge, I. S. (1990) *J. Cell Biol.* **10,** 283–294
22. Zervos, A., Gyuris, J., and Brent, R. (1993) *Cell* **72,** 223–232
23. Gyuris, J., Golemis, E., Chertkov, H., and Brent, R. (1993) *Cell* **75,** 791–803
24. Seino, S., Seino, M., Nishi, S., and Bell, G. I. (1989) *Proc. Natl. Acad. Sci. U. S. A.* **86,** 114–118
25. Knighton, D. R., Zheng, J., Ten Eyck, L. F., Ashford, V. A., Zuong, H.-H., Taylor, S. S., and Sowadski, J. M. (1991) *Science* **253,** 407–414
26. Schiest, R. H., and Gietz, R. D. (1986) *Curr. Genet.* **16,** 339–346
27. Strathmann, M., Hamilton, B. A., Mayeda, C. A., Simon, M. I, Meyerowitz, E. M., and Palazzolo, M. J. (1991) *Proc. Natl. Acad. Sci. U. S. A.* **88,** 1247–1250
28. Wiley, H. S. (1988) *J. Cell Biol.* **107,** 801–810
29. Backer, J. M., Shoelson, S. E., Haring, E., and White, M. (1991) *J. Cell Biol.* **115,** 1535–1545
30. Ullrich, A., Gray, A., Tam, A. W., Yang-Feng, T., Tsubokawa, M., Collins, C., Henzel, W., Le Bon, T., Kathuria, S., Chen, E., Jacobs, S., Srancke, U., Ramachandran, J., and Fujita-Yamaguchi, Y. (1986) *EMBO J.* **5,** 2503–2512

31. Chen, W. S., Lazar, C. S., Lund, K. A., Welsh, J. B., Chang, C.-P., Walton, G. M., Der, C. J., Wiley, H. S., Gill, G. N., and Rosenfeld, M. G. (1989) *Cell* **59,** 33–43
32. Pérez-Alvarado, G. C., Miles, C., Michelsen, J. W., Louis, H. A., Winge, D. R., Beckerle, M. C., and Summers, M. F. (1994) *Nature Structural Biol.* **1,** 388–397
33. Cohen, B., McGuffin, M. E., Pfeifle, C., Segal, D., and Cohen, S. M. (1992) *Genes & Dev.* **6,** 715–729
34. Xu, Y., Baldassare, M., Fisher, P., Rathbun, G., Oltz, E. M., Yancopoulous, G. D., Jessell, T. M., and Alt, F. W. (1993) *Proc. Natl. Acad. Sci. U. S. A.* **90,** 227–231
35. Barnes, J. D., Crosby, J. L., Jones, C. M., Wright, C. V. E., and Hogan, B. L. M. (1994) *Dev. & Biol.* **161,** 168–178
36. Pawson, T., and Gish, G. D. (1992) *Cell* **71,** 359–362
37. German, M. S., Wang, J. R., Chadwick, R. B., and Rutter, W. J. (1992) *Genes & Dev.* **6,** 2165–2176
38. Songyang, Z., Shoelson, S. E., Chaudhuri, M., Gish, G., Pawson, T., Haser, W. G., King, F., Roberts, T., Ratnofsky, S., Lechleider, R. J., Neel, B. G., Birge, R. B., Fajardo, J. E., Chou, M. M., Hanafusa, H., Schaffhausen, B., and Cantley, L. C. (1993) *Cell* **72,** 767–778

Communication

THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 271, No. 22, Issue of May 31, pp. 12691–12694, 1996

Mitogenic Signaling by Ret/ptc2 Requires Association with Enigma via a LIM Domain*

(Received for publication, March 8, 1996, and in revised form, March 29, 1996)

Kyle Durick‡§, Rui-Yun Wu¶‖, Gordon N. Gill, and Susan S. Taylor‡ ‡‡**

*From the Departments of ‡Chemistry and Biochemistry, ¶Biology, and **Medicine, University of California, San Diego, La Jolla, California 92093-0654*

The *ret/ptc2* papillary thyroid cancer oncogene, an oncogenic form of the c-Ret receptor tyrosine kinase, is the product of a somatic crossover event fusing the dimerization domain of the type Iα regulatory subunit of cyclic AMP-dependent protein kinase (RI) with the tyrosine kinase domain of c-Ret. Mitogenic activity of Ret/ptc2 required dimerization via the N terminus of RI and a tyrosine residue located C-terminal to the kinase core of Ret, Tyr-586 (Durick, K., Yao, V. J., Borrello, M. G., Bongarzone, I., Pierotti, M. A. and Taylor, S. S. (1995) *J. Biol. Chem.* 270, 24642–24645). Using the yeast two-hybrid system, Ret/ptc2 binding proteins were identified, and the sites of interaction with Ret/ptc2 were mapped. The SH2 domains of phospholipase Cγ and Grb10 were both identified, and binding depended on phosphorylation of Tyr-539 and Tyr-429, respectively. These interactions, however, were not required for mitogenic signaling. The second of the three LIM domains in Enigma (Wu, R. Y., and Gill, G. N. (1994) *J. Biol. Chem.* 269, 25085–25090) was also identified as a Ret/ptc2 binding domain. Enigma, a 455-residue protein, was discovered based on its interaction with the insulin receptor through the C-terminal LIM domain. Although the association with Enigma required Tyr-586 of Ret/ptc2, the interaction was phosphorylation-independent. In contrast to the SH2 interactions, disruption of the interaction with Enigma abolished Ret/ptc2 mitogenic signaling, suggesting that LIM domain recognition of an unphosphorylated tyrosine-based motif is required for Ret signal transduction.

The *c-ret* proto-oncogene encodes a receptor tyrosine kinase with a cadherin-like extracellular domain (1). Mutations of *c-ret* are responsible for two distinct classes of genetic disease. Germline loss of function mutations in *c-ret* result in the developmental disorder Hirschsprung's disease (2, 3), while activating mutations result in the multiple endocrine neoplasia family of inherited cancers (4–7). In addition to the germline alterations, somatic mutational events lead to constitutively active forms of *c-ret*, and these are found in nearly half of all papillary type thyroid carcinomas (8).

Chromosome translocations or inversions in papillary thyroid carcinoma (ptc)[1] give rise to various fusion proteins where the C-terminal tyrosine kinase domain of c-Ret is fused to an N-terminal portion of another gene product. One of the resultant transforming proteins, observed in multiple independent cases of ptc, was the product of a crossover between the type Iα regulatory subunit of cyclic AMP-dependent protein kinase (RI) gene with *c-ret* (9). This protein, Ret/ptc2, is 596 residues in length and contains the N-terminal two-thirds of RI followed by the entire tyrosine kinase domain of c-Ret (10). Using a microinjection-based assay for mitogenic activity, we previously showed that the N-terminal dimerization domain of RI was essential for constitutive activation of Ret/ptc2 (11).

In the absence of a known ligand, studies of Ret signaling have been done using activated forms of the Ret tyrosine kinase, like Ret/ptc2, or a chimeric epidermal growth factor/Ret kinase receptor. It was suggested from work with the EGFR/Ret chimera that Ret couples to a novel mitogenic signaling pathway because, while growth stimulatory effects were as strong as those of the EGF or platelet-derived growth factor receptors, stimulation of mitogen-activated protein kinases and PLCγ by Ret was, in comparison, very weak (12). Work with Ret/ptc2 indicated that both intrinsic protein tyrosine kinase activity and a tyrosine residue (Tyr-586) located outside the kinase core were absolutely required for Ret/ptc2-induced mitogenesis (11).

A yeast two-hybrid screen was used to identify the presumed Src homology 2 (SH2) or phosphotyrosine binding domain that interacts with Tyr-586 of Ret/ptc2. The SH2 domains of PLCγ and Grb10 were both found to interact with Ret, but neither bound at Tyr-586. Mutations in Ret/ptc2, which interfered with these SH2 interactions, had no effect on the mitogenic activity. A protein that interacted with Tyr-586 was identified as Enigma (13). This interaction was found to be highly specific, mediated by the second of the three LIM domains of Enigma and independent of Ret phosphorylation or activity. Disruption of the interaction of Ret/ptc2 with Enigma, either by mutation of Tyr-586 in Ret/ptc2 or by co-expression with a dominant negative form of Enigma, abolished the mitogenic activity of Ret/ptc2.

EXPERIMENTAL PROCEDURES

Two-hybrid Screen—A yeast two-hybrid screen was performed by the methods of Vojtek *et al.* (14), with reagents from Stan Hollenberg. Ret/ptc2 cDNA was subcloned into the LexA-fusion vector pBTM116 and coexpressed in the L40 strain of *Saccharomyces cerevisiae* with an embryonic mouse random primed cDNA library. From approximately two million co-transformants, seven interacted specifically with the Ret portion of Ret/ptc2. The cDNA inserts of these were sequenced by the dideoxy method (15), and sequences obtained were compared with the contents of GenBank using the BLAST program through the NIH/NCBI

* This research was supported in part by United States Army Grant AIBS1762 (to S. S. T.) and National Institutes of Health Grant DK13149 (to G. N. G.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

§ Supported by the Markey Charitable Trust as a Fellow and currently supported by National Institutes of Health Training Grant NCI T32 CA09523.

‖ Supported by United States Army Breast Cancer Research Fellowship DAMD 17-94-J-4124.

‡‡ To whom correspondence should be addressed: 9500 Gilman Dr., 0654, University of California, San Diego, La Jolla, CA 92093-0654. Tel.: 619-534-3677; Fax: 619-534-8193; E-mail: staylor@ucsd.edu.

[1] The abbreviations used are: ptc, papillary thyroid carcinoma; RI, type Iα regulatory subunit of cyclic AMP-dependent protein kinase; PLCγ, phospholipase C-γ; GST, glutathione *S*-transferase; EGF, epidermal growth factor; EGFR, epidermal growth factor receptor; SH2, Src homology 2.

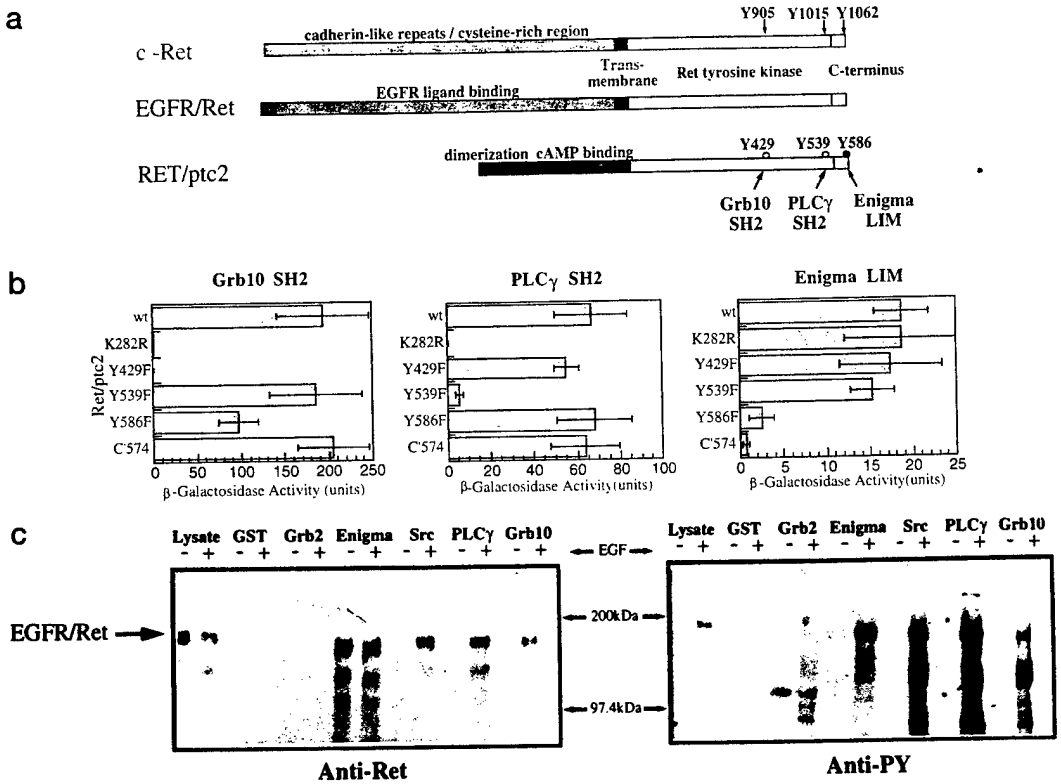


FIG. 1. **Requirements for interaction of Grb10, PLCγ, and Enigma with Ret.** *a*, schematic representation showing the location of tyrosines that are required for interaction with the SH2 domains of Grb10 and PLCγ and the second LIM domain of Enigma. The corresponding residues in c-Ret are indicated, and the EGFR/Ret chimera used in *panel c* is shown. *b*, analysis of mutant Ret/ptc2 interactions using the yeast two-hybrid system. The indicated mutant of Ret/ptc2 and Grb10 SH2, PLCγ SH2, or LIM 2/3 of Enigma were co-expressed in yeast, and interactions were measured using solution assays of the reporter β-galactosidase. Y586F also contained R588T. The mean ± S.D. (n = 6) for each interacting pair is shown. *c, in vitro* interactions between GST-fusion proteins and the EGFR/Ret chimeric receptor. Clonal NIH3T3 cells expressing an EGFR/Ret chimeric receptor were treated (+) or not treated (−) with EGF before lysis. Western blots of EGFR/Ret that bound to the indicated GST-fusion proteins are shown. Gels were run in parallel, blotted to polyvinylidene difluoride membranes, and probed with anti-Ret (*left*) or anti-phosphotyrosine antibodies (*right*).

server on the World Wide Web. Three library vectors encoded the following mouse sequences: the C-terminal 155 residues of Grb10 (16); 156 residues that share 97% identity with residues 537–693 of rat PLCγ2 (17); and 131 residues with 95% identity to the C-terminal 131 residues of human Enigma containing all of LIM2 and LIM3 (13).

β-Galactosidase Assay—Two hybrid transformants were assayed for β-galactosidase activity by solution assay (18). Units of activity were calculated as: activity = $1750(A_{420})$/((time in min)(volume of culture in assay)(A_{600} of culture)).

GST-Fusion Affinity Precipitation—Two-hybrid results were verified using a stably transfected NIH3T3 cell line expressing an EGFR/Ret chimeric protein (12). These cells were treated with 100 nM EGF for 10 min before resuspension in lysis buffer (50 mM HEPES, pH 7.4, 150 mM NaCl, 5 mM KCl, 1 mM $CaCl_2$, 1 mM $MgSO_4$, 10% glycerol, 1% Triton, 1 mM benzamidine, 1 mM tosylphenylalanyl chloromethyl ketone, 1 mM N^{α}-*p*-tosyl-L-lysine chloromethyl ketone, 1 mM phenylmethylsulfonyl fluoride, 1 mM $NaVO_4$). Cleared lysates were incubated for 2 h with 2 μg of GST-fusion protein bound to glutathione-agarose beads in a total volume of 300 μl. The beads were washed four times with lysis buffer, resuspended in SDS-polyacrylamide gel electrophoresis sample buffer, boiled, and run on 7% gels. Proteins were transferred to polyvinylidene difluoride membranes and probed with either rabbit anti-Ret (11) (1:100,000) or anti-phosphotyrosine (1:2500, Transduction Laboratories) antibodies. The GST-fusion proteins used were bacterially expressed from pGEX vectors coding for the following: GST, empty vector; GST-Grb2, murine Grb2 SH2 domain; GST-Enigma, human Enigma LIM2 domain (residues 334–394); GST-Src, murine v-Src SH2 domain; GST-PLCγ, murine PLCγ SH2 domain 1 obtained from the two-hybrid screen; GST-Grb10, murine Grb10 SH2 domain obtained from the two-hybrid screen.

Microinjection Mitogenic Activity Assay—The development of this assay is described in detail elsewhere (11). Briefly, mouse 10T1/2 fibroblasts were plated on glass coverslips and grown to 70% confluence in Dulbecco's modified Eagle's medium containing 10% fetal bovine serum. The coverslips were then transferred to Dulbecco's modified Eagle's medium containing 0.05% calf serum. After 24 h of serum starvation, the cells were injected into their nuclei with solutions of injection buffer (20 mM Tris, pH 7.2, 2 mM $MgCl_2$, 0.1 mM EDTA, 20 mM NaCl) containing 100 μg/ml Ret/ptc2 expression plasmid DNA and 8 mg/ml rabbit IgG (Sigma). For co-injection experiments, 200 μg/ml of a second expression plasmid was also present. All microinjection experiments were performed using an automatic micromanipulator (Eppendorf, Fremont, CA), with glass needles pulled on a vertical pipette puller (Kopf, Tujunga, CA). Entry into S-phase was assessed through incorporation of the thymidine analog 5-bromodeoxyuridine and its subsequent detection by immunostaining. Injected cells were identified by immunostaining of the rabbit IgG injection marker.

RESULTS

Yeast Two-hybrid Interactions—To search for proteins that interact with the Ret/ptc2 oncogenic protein, a mouse random-primed cDNA library was screened using a yeast two-hybrid system (14). Three sequences isolated from the library by interaction with Ret/ptc2 matched the SH2 domain of Grb10, the first SH2 domain of PLCγ, and a C-terminal fragment of Enigma that contained LIM domains 2 and 3 (LIM2/3). The interaction of Ret with PLCγ and Grb10 has been observed previously (12, 19). Using the two-hybrid system, the interactions of these proteins with Ret were characterized.

Mutants of Ret/ptc2 were prepared and interactions were quantitated by β-galactosidase activity (Fig. 1*b*). The SH2 domains failed to interact with a kinase-inactive mutant of Ret/ptc2 (K282R), indicating that these interactions depended on autophosphorylated tyrosine residues. In contrast to the SH2 domains, binding of the LIM2/3 was not diminished in the kinase-inactive mutant. By testing for interaction between the SH2 domains and various Tyr to Phe mutants of Ret/ptc2,

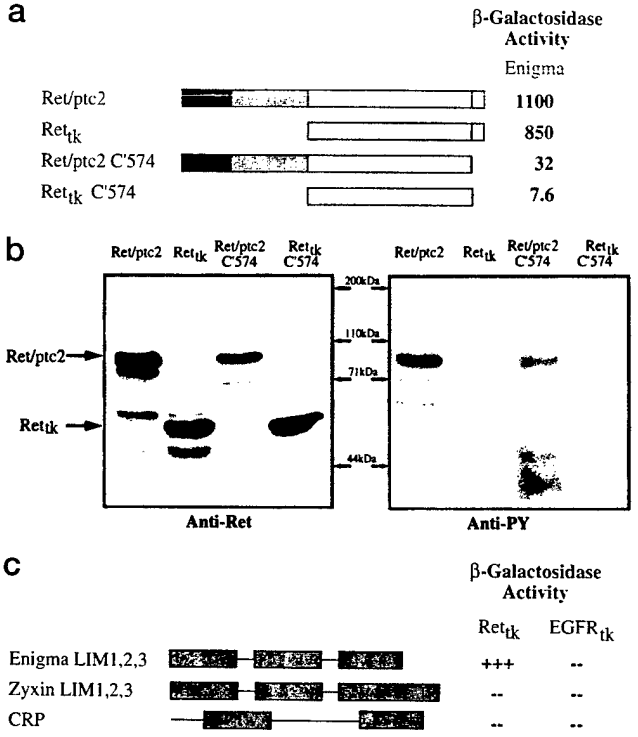


FIG. 2. **Requirements for the association between Ret/ptc2 and Enigma.** *a*, mapping of binding determinants in Ret/ptc2. Various fragments of Ret/ptc2, shown schematically, were used in the yeast two-hybrid system to measure interaction with the product of a plasmid expressing full-length Enigma. β-Galactosidase activity of transformants was measured by solution assay, and values shown are averages of duplicate assays. The same pattern was observed in results from three separate experiments with full-length Enigma, the C-terminal half containing the LIM domains, or LIM2 alone. *b*, tyrosine autophosphorylation of Ret/ptc2 mutants. Lysates of yeast transformed with plasmids expressing fragments shown in *panel a* were run on 10% SDS-polyacrylamide gel electrophoresis, blotted to nitrocellulose, and probed as in Fig. 1*c*. *Anti-PY*, anti-phosphotyrosine. *c*, specificity of the LIM domains of Enigma for Ret/ptc2. pJG4–5 plasmids coding for full-length CRP, the three LIM domains of Zyxin (residues 339–542), and the three LIM domains of Enigma (residues 275–455) were co-transformed into yeast with either Ret$_{tk}$, the pEG202 plasmid expressing the Ret kinase with an intact C terminus, or EGFR$_{tk}$, a pEG202 construct containing the intracellular domain of EGFR. β-Galactosidase activity was assessed by streaking transformants on 5-bromo-4-chloro-3-indoyl β-D-galactoside plates. +++, denotes dark blue after 12 h of incubation compared with white (– –) that was equivalent to background after 24 h.

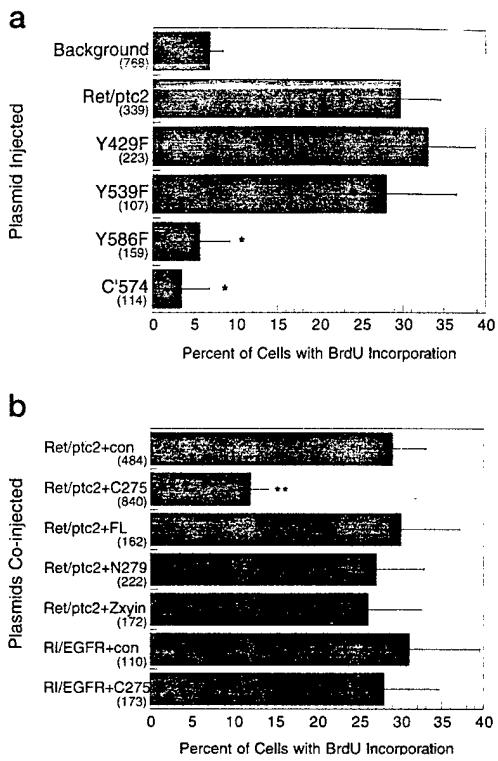


FIG. 3. **Effect of mutations in Ret that disrupt association with Enigma and of co-expression with the LIM domains of Enigma on mitogenic activity of RET/ptc2.** *a*, effects of mutations in Ret/ptc2. Serum-starved mouse fibroblasts (10T1/2) were microinjected with plasmids expressing either wild-type Ret/ptc2 or various mutants and then assessed for entry into S-phase by immunofluorescent detection of 5-bromodeoxyuridine (*BrdU*) incorporation. The fraction of injected cells positive for 5-bromodeoxyuridine incorporation is shown with *error bars* displaying the 95% confidence interval calculated using the standard error of proportion. The *numbers* in *parentheses* are the total number of injected cells. Plasmids were injected at a concentration of 100 μg/ml. *Asterisk* denotes a highly significant difference between cells injected with Ret/ptc2 or Y586F and Ret/ptc2 or C'574 ($p < 0.001$). *b*, effects of co-expression of Enigma and its fragments. Serum-starved fibroblasts were microinjected with a mixture of two of the following expression plasmids: Ret/ptc2 or RI/EGFR, a construct analogous to Ret/ptc2 with the EGFR intracellular domain in place of the Ret kinase, plus either *con*, control empty plasmid; *C275*, the C-terminal 275 residues of Enigma containing LIM domains 1, 2, and 3; *FL*, full-length Enigma; *N279*, N-terminal 279 residues of Enigma lacking LIM domains; or *Zyxin*, LIM domains from Zyxin, residues 339–452. In each case Ret/ptc2 and RI/EGFR constructs were injected at 100 μg/ml, while the other constructs were present at 200 μg/ml. *Double asterisk* denotes a highly significant difference between cells injected with Ret/ptc2+con or Ret/ptc2+C275 ($p < 0.001$).

residues required for Grb10 and PLCγ binding were identified as Tyr-429 and Tyr-539, respectively. Both the Grb10 and PLCγ SH2 domains bound to the Y586F mutant, but binding to the Enigma LIM domains was eliminated by this mutation and by a mutation in which the C-terminal 23 residues of Ret were deleted (C'574).

In Vitro Binding to Ret—To verify the two-hybrid results, affinity precipitation experiments were performed. The binding domains from PLCγ, Grb10, and Enigma were expressed in *Escherichia coli* as GST-fusion proteins. The three GST-fusion proteins were incubated with lysates of NIH3T3 cells expressing the EGFR/Ret chimeric receptor (Fig. 1*a*), where EGF-dependent activation of EGFR/Ret has been characterized (12). In each case binding to EGFR/Ret from lysates of EGF-treated cells was observed (Fig. 1*c*). Results shown were using a GST-fusion protein of only LIM2 from Enigma, because Enigma binding to Ret was determined to be mediated by LIM2.[2] GST-fusion proteins with the SH2 domains of Grb2 and v-Src were also expressed and tested for *in vitro* binding. Neither GST alone nor GST-Grb2 bound to the EGFR/Ret chimera. The SH2 domain of v-Src, however, did interact with EGFR/Ret, and interaction with all three GST-SH2 domains required EGF-stimulated receptor autophosphorylation. In contrast, interaction with GST-LIM2 of Enigma did not require receptor autophosphorylation. Grb2 served as a negative control because Ret has two splice isoforms (10). The long form binds to Grb2 (20) while the short form, used in all of the constructs described here, does not contain the Grb2 consensus site. Both isoforms of Ret are mitogenic (12).

Characterization of the Ret-Enigma Interaction—Because the LIM2 domain of Enigma bound at a site crucial for the mitogenic activity of Ret, this interaction was investigated further. Using an inducible two-hybrid system (21, 22), where higher expression levels were achieved, it was possible to observe the phosphorylation state of Ret using anti-phosphoty-

[2] Wu, R.-Y., Durick, R., Songyang, Z., Cantley, L. C., Taylor, S. S., and Gill, G. N. (1996) *J. Biol. Chem.*, in press.

rosine antibodies. Enigma bound to both Ret/ptc2 and to the Ret tyrosine kinase alone (Ret_{tk}) but did not bind to either when the C-terminal 23 residues were deleted (C'574, Fig. 2*a*). The interaction was not dependent upon the phosphorylation state of Ret because Ret_{tk} was not phosphorylated on tyrosine in the absence of the dimerization domain of RI, whereas Ret/ptc2 underwent tyrosine autophosphorylation (Fig. 2*b*, *lanes* Ret_{tk} and *Ret/ptc2*).

Based on optical densitometry, the extent of tyrosine phosphorylation of Ret/ptc2 and the C-terminal deletion Ret/ptc2 C'574 were equivalent. Because deletion of the C terminus of Ret/ptc2 did not decrease the phosphotyrosine content by a detectable amount, Tyr-586, the only tyrosine in the C terminus of this isoform of Ret, does not appear to be a major site of autophosphorylation. Results using both the EGFR/Ret chimera and Ret/ptc2 thus indicated that interaction with Enigma required the C terminus containing Tyr-586 but was independent of tyrosine autophosphorylation of Ret.

The interaction between LIM2 of Enigma and Ret was specific because LIM domains from other proteins failed to interact with Ret (Fig. 2*c*). Another tyrosine kinase, EGFR, failed to interact with the LIM domains of Enigma either in the two-hybrid system (Fig. 2*c*) or in GST-fusion binding reactions (data not shown).

Effect of Enigma on Ret/ptc2 Mitogenic Activity—The functional significance of the association between Ret and Enigma was investigated *in vivo* using microinjection. Various Ret/ptc2 expression constructs were injected into nuclei of serum-starved fibroblasts, and the capacity of these constructs to induce DNA synthesis was assayed by monitoring incorporation of the thymidine analog 5-bromodeoxyuridine. Mutations in Ret/ptc2 that blocked association with Grb10 (Y429F) or PLCγ (Y539F) had no significant effect on mitogenic activity, while mutation of Tyr-586 to Phe or deletion of the C terminus of Ret/ptc2 completely blocked the ability of Ret/ptc2 to induce DNA synthesis (Fig. 3*a*).

The strong correlation between mutants that failed to bind Enigma and loss of mitogenic activity suggested that Enigma was either required for the mitogenic signaling of Ret/ptc2 or that some other protein, which was not detected in the two-hybrid screen, also binds at Tyr-586. To discriminate between these two possibilities, co-injection experiments were performed to attempt to block the Ret/ptc2 mitogenic signal. Co-injection of Ret/ptc2 with a plasmid that expressed the three LIM domains of Enigma (C275, Fig. 3*b*) blocked Ret-induced DNA synthesis, while co-injection with full-length Enigma had no effect. These results support the conclusion that Ret/ptc2 requires Enigma for mitogenic signaling. If Enigma was simply competing for binding with some other signaling protein, then both full-length Enigma and the LIM domains alone should block signaling, given that both interact with Ret with an equivalent affinity.[2] The inhibition of mitogenesis was specific to the LIM domains of Enigma because the LIM domains of Zyxin (23) were without effect. It was also specific for Ret because the LIM domains of Enigma did not block the mitogenic activity of the EGFR tyrosine kinase analog of Ret/ptc2 (RI/EGFR) previously shown to have mitogenic activity in this assay (11).

DISCUSSION

LIM domains contain approximately 50 amino acids, bind two atoms of Zn^{2+}, and are found in a variety of homeodomain proteins (24), cytoskeleton-associated proteins (23, 25), protein kinases, and proteins of unknown function (26). Enigma was originally discovered as a protein that binds to exon 16 of the insulin receptor (13) at a tyrosine-based sequence important for receptor internalization, and that interaction is through the C-terminal LIM domain of Enigma, LIM3. In the present study, Enigma bound to Ret via the LIM2 domain to a sequence required for mitogenic signaling. LIM3 of Enigma is highly specific for the Tyr-based motif in the insulin receptor whereas LIM2 is highly specific for the Tyr-586-based motif in Ret.[2] The N-terminal portion of Enigma is required for mitogenic signaling because co-injection of only the LIM domains with Ret/ptc2 ablated the mitogenic signal, while co-injection with full-length Enigma did not.

As shown here, Enigma is required for Ret/ptc2 mitogenic signaling while previous results established that Ret tyrosine kinase activity was also required (11). Kinase activity, however, is not required for recruitment of Enigma to Ret/ptc2. Enigma binds to a tyrosine-containing sequence in an activation-independent manner, and this is clearly different from SH2 or phosphotyrosine binding domain interactions. Enigma might either become phosphorylated upon Ret activation or serve to localize Ret to a subcellular position required for kinase-mediated signaling. In either case, these results define a novel mechanism for mitogenic signaling.

Acknowledgments—We thank P. Paolo Di Fiore, European Institute of Oncology, Milan, Italy, for cells expressing the EGFR/Ret chimera, M. Pierotti, Institute Nationale Tumori, Milan, Italy, for the Ret/ptc2 cDNA, and Tony Hunter and David Schlaepfer, Salk Institute, La Jolla, CA, for the pGEX-Grb2 and pGEX-Src constructs.

REFERENCES

1. Takahashi, M., Buma, Y., and Hiai, H. (1989) *Oncogene* **4,** 805–806
2. Romeo, G., Ronchetto, P., Luo, Y., Barone, V., Seri, M., Ceccherini, I., Pasini, B., Bocciardi, R., Lerone, M., Kaariainen, H., and Martucciello, G. (1994) *Nature* **367,** 377–378
3. Edery, P., Lyonnet, S., Mulligan, L. M., Pelet, A., Dow, E., Abel, L., Holder, S., Nihoul-Fekete, C., Ponder, B. A., and Munnich, A. (1994) *Nature* **367,** 378–380
4. Mulligan, L. M., Kwok, J. B. J., Healey, C. S., Elsdon, M. J., Eng, C., Gardner, E., Love, D. R., Mole, S. E., Moore, J. K., Papi, L., Ponder, M. A., Telenius, H., Tunnacliffe, A., and Ponder, B. A. J. (1993) *Nature* **363,** 458–460
5. Mulligan, L. M., Eng, C., Healey, C. S., Clayton, D., Kwok, J. B., Gardner, E., Ponder, M. A., Frilling, A., Jackson, C. E., Lehnert, H., Neumann, H., Thibodeau, S., and Ponder, B. (1994) *Nat. Genet.* **6,** 70–74
6. Eng, C., Smith, D. P., Mulligan, L. M., Nagai, M. A., Healey, C. S., Ponder, M. A., Gardner, E., Scheumann, G. F., Jackson, C. E., Tunnacliffe, A., and Ponder, B. (1994) *Hum. Mol. Genet.* **3,** 237–241
7. Hofstra, R. M., Landsvater, R. M., Ceccherini, I., Stulp, R. P., Stelwagen, T., Luo, Y., Pasini, B., Hoppener, J. W. M., Ploos van Amstel, H. K., Romeo, G., Lips, C. J. M., and Buys, C. C. C. M. (1994) *Nature* **367,** 375–376
8. Bongarzone, I., Butti, M. G., Coronelli, S., Borrello, M. G., Santoro, M., Mondellini, P., Pilotti, S., Fusco, A., Della Porta, G., and Pierotti, M. A. (1994) *Cancer Res.* **54,** 2979–2985
9. Lanzi, C., Borrello, M. G., Bongarzone, I., Migliazza, A., Fusco, A., Grieco, M., Santoro, M., Gambetta, R. A., Zunino, F., Della Porta, G., and Pierotti, M. A. (1992) *Oncogene* **7,** 2189–2194
10. Bongarzone, I., Monzini, N., Borrello, M. G., Carcano, C., Ferraresi, G., Arighi, E., Mondellini, P., Della Porta, G., and Pierotti, M. A. (1993) *Mol. Cell. Biol.* **13,** 358–366
11. Durick, K., Yao, V. J., Borrello, M. G., Bongarzone, I., Pierotti, M. A., and Taylor, S. S. (1995) *J. Biol. Chem.* **270,** 24642–24645
12. Santoro, M., Wong, W. T., Aroca, P., Santos, E., Matoskova, B., Grieco, M., Fusco, A., and Di Fiore, P. P. (1994) *Mol. Cell. Biol.* **14,** 663–675
13. Wu, R. Y., and Gill, G. N. (1994) *J. Biol. Chem.* **269,** 25085–25090
14. Vojtek, A. B., Hollenberg, S. M., and Cooper, J. A. (1993) *Cell* **74,** 205–214
15. Sanger, F., Nicklen, S., and Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U. S. A.* **74,** 5463–5467
16. Ooi, J., Yajnik, V., Immanuel, D., Gordon, M., Moskow, J. J., Buchberg, A. M., and Margolis, B. (1995) *Oncogene* **10,** 1621–1630
17. Emori, Y., Homma, Y., Sorimachi, H., Kawasaki, H., Nakanishi, O., Suzuki, K., and Takenawa, T. (1989) *J. Biol. Chem.* **264,** 21885–21890
18. Ausubel, F., Brent, R., Kingston, R., Moore, D., Seidman, J., Smith, J., and Struhl, K. (1994) *Current Protocols in Molecular Biology,* John Wiley & Sons, Inc., New York
19. Pandey, A., Duan, H., Di Fiore, P. P., and Dixit, V. M. (1995) *J. Biol. Chem.* **270,** 21461–21463
20. Borrello, M. G., Pelicci, G., Arighi, E., De Filippis, L., Greco, A., Bongarzone, I., Rizzetti, M., Pelicci, P. G., and Pierotti, M. A. (1994) *Oncogene* **9,** 1661–1668
21. Zervos, A., Gyuris, J., and Brent, R. (1993) *Cell* **72,** 223–232
22. Gyuris, J., Golemis, E., Chertkov, H., and Brent, R. (1993) *Cell* **75,** 791–803
23. Sadler, I., Crawford, A. W., Michelssen, J. W., and Beckerle, M. C. (1992) *J. Cell Biol.* **119,** 1573–1587
24. Taira, M. M., Jamrich, M., Good, P. J., and Dawid, I. B. (1992) *Genes & Dev.* **6,** 356–366
25. Turner, C. E. (1994) *Bioessays* **16,** 47–52
26. Gill, G. N. (1995) *Structure* **3,** 1285–1289

THE JOURNAL OF BIOLOGICAL CHEMISTRY

Vol. 271, No. 27, Issue of July 5, pp. 15934–15941, 1996
Printed in U.S.A.

Specificity of LIM Domain Interactions with Receptor Tyrosine Kinases*

(Received for publication, January 22, 1996, and in revised form, April 11, 1996)

Rui-yun Wu‡§, Kyle Durick¶‖, Zhou Songyang‡‡, Lewis C. Cantley‡‡, Susan S. Taylor¶, and Gordon N. Gill§§¶¶

From the Departments of ‡Biology, §§Medicine, and ¶Chemistry and Biochemistry, University of California San Diego, La Jolla, California 92093-0650 and ‡‡Harvard Medical School, Cambridge, Massachusetts 02115

LIM domains, Cys-rich motifs containing approximately 50 amino acids found in a variety of proteins, are proposed to direct protein-protein interactions. To identify structural targets recognized by LIM domains, we have utilized random peptide library selection, the yeast two-hybrid system, and glutathione *S*-transferase fusions. Enigma contains three LIM domains within its carboxyl terminus and LIM3 of Enigma specifically recognizes active but not mutant endocytic codes of the insulin receptor (InsR) (Wu, R. Y., and Gill, G. N. (1994) *J. Biol. Chem.* 269, 25085–25090). Interaction of two random peptide libraries with glutathione *S*-transferase-LIM3 of Enigma indicated specific binding to Gly-Pro-Hyd-Gly-Pro-Hyd-Tyr-Ala corresponding to the major endocytic code of InsR. Peptide competition demonstrated that both Pro and Tyr residues were required for specific interaction of InsR with Enigma. In contrast to LIM3 of Enigma binding to InsR, LIM2 of Enigma associated specifically with the receptor tyrosine kinase, Ret. Ret was specific for LIM2 of Enigma and did not bind other LIM domains tested. Mutational analysis indicated that the residues responsible for binding to Enigma were localized to the carboxyl-terminal 61 amino acids of Ret. A peptide corresponding to the carboxyl-terminal 20 amino acids of Ret dissociated Enigma and Ret complexes, while a mutant that changed Asn-Lys-Leu-Tyr in the peptide to Ala-Lys-Leu-Ala or a peptide corresponding to exon16 of InsR failed to disrupt the complexes, indicating the Asn-Lys-Leu-Tyr sequence of Ret is essential to the recognition motif for LIM2 of Enigma. We conclude that LIM domains of Enigma recognize tyrosine-containing motifs with specificity residing in both the LIM domains and in the target structures.

LIM domains contain about 50 amino acids and have a cysteine-rich sequence: $CX_2CX_{16-23}CX_2HX_2CX_2CX_2CX_{16-23}CX_2C$ (1). They bind two atoms of Zn^{2+} with coordination being S_4, S_3N_1, S_3O_1, or $S_2N_1O_1$ (2). The NMR structure of the carboxyl-terminal LIM domain of chicken cysteine-rich protein (CRP)[1] revealed that the two Zn^{2+} binding modules are located at the ends of a hydrophobic core composed of antiparallel β sheets (3). LIM domains were initially recognized in the primary sequences of the homeodomain proteins Lin 11 (4), Isl-1 (5), and Mec-3 (6) and have subsequently been identified in a variety of homeodomain proteins (7, 8), in cytoskeleton-associated proteins (9, 10), in LIM domain-only proteins (11–13), in protein kinases (14), and in proteins of undefined function (8).

Most LIM proteins contain more than one LIM domain. The sequence of an individual LIM domain is, in general, more closely related to the same LIM domain in analogous proteins from other species than to other LIM domains within the same protein (8). Although the NMR structure of LIM2 of CRP resembles the DNA binding domain of the GATA-1 transcription factor (3), most available evidence indicates that LIM domains function in protein-protein rather than protein-DNA interactions.

Two structural targets for LIM domains have been identified. Using gel overlay techniques, Schmeichel and Beckerle (15) found that the LIM domains of zyxin interacted with the LIM-only protein CRP. Specificity was evident from the observation that LIM1 but not LIM2 or LIM3 of zyxin-bound CRP. Feurstein, *et al.* (16) also found evidence for LIM-LIM interactions involving CRP but did not observe specificity for the LIM domain. The carboxyl-terminal LIM domain of the cytoplasmic protein Enigma was found to specifically interact with exon 16 of the insulin receptor (InsR) (17). Mutations in exon 16 that disrupted the major endocytic code and ligand-induced endocytosis of InsR (18) also disrupted interaction with Enigma. The endocytic code of InsR, like that of many receptors (19), consists of 4–6 amino acids that form a tyrosine-containing tight turn (20). A generalized tight-turn motif, which functioned in endocytosis of mutant EGFR (21), and which contained two copies of an Asn-Asn-Ala-Tyr-Phe motif interacted with a wider range of LIM domains, suggesting that specific Tyr-based tight turns would provide interaction targets for specific LIM domains. There is functional evidence for LIM domain interactions with a variety of transcription factors (22–24), suggesting that additional target specificities exist.

We recently found that in addition to InsR, Enigma specifically interacted with the receptor tyrosine kinase Ret (25). Given that Enigma, which contains three LIM domains at its carboxyl terminus,[2] was found to interact with two receptor tyrosine kinases, determining the molecular basis of this recognition became important. We found that LIM2 of Enigma

* These studies were supported in part by National Institutes of Health Grant DK 13149 (to G. N. G.) and Department of the Army Grant DAMD 17-94-J-4120 (to S. S. T.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

§ Supported by United States Army Breast Cancer Research Predoctoral Fellowship, DAMD 17-94-J-4124.

‖ Supported by National Institutes of Health Training Grant T32CA 02523 and in part as a fellow of the Markey Charitable Trust.

¶¶ To whom correspondence should be addressed: University of California San Diego, 9500 Gilman Drive 0650, La Jolla, CA 92093-0650. Tel.: 619-534-4310; Fax: 619-534-8193; E-mail: ggill@ucsd.edu.

[1] The abbreviations used are: CRP, cysteine-rich protein; InsR, insulin receptor; Ret, receptor tyrosine kinase; EGF, epidermal growth factor; EGFR, epidermal growth factor receptor; GST, glutathione *S*-transferase; HA, hemagglutinin; PTB, phosphotyrosine binding.

[2] The LIM domains of Enigma are numbered from amino- to carboxyl-terminal location in the protein. LIM3 equals LIM2 in Wu and Gill (17).

specifically recognized Ret whereas LIM3 of Enigma specifically recognized InsR. Detailed analysis of the target sites indicated both LIM2 and LIM3 recognize Tyr-containing motifs located outside the tyrosine kinase cores of Ret and InsR. Individual LIM domains thus have the ability to distinguish between Tyr-based motifs providing a mechanism for specificity in both the LIM domain and in the target. This is of special interest given the requirement of the target of LIM2 in Ret for mitogenic signaling and of the target of LIM3 in InsR for endocytosis.

EXPERIMENTAL PROCEDURES

Materials—The *Caenorhabditis elegans* Mec-3 cDNA was obtained from Dr. Serge Lichsteiner, University of California Berkeley, Berkeley, CA; the zyxin and CRP cDNAs were obtained from Dr. Mary C. Beckerle, University of Utah, Salt Lake City, Utah; the cDNAs of Isl-1 and Xlm-1 were obtained from Dr. Mark Montminy, The Salk Institute, La Jolla, CA; and the cDNA encoding the LIM domains of paxillin were obtained from Dr. Michael Brown, State University of New York, Syracuse, NY. The Ret/ptc2 clone was a gift from Dr. M. Pierotti, Institute Nationale Tumori, Milan, Italy. NIH3T3 cells overexpressing the EGFR/Ret chimera were a gift of Dr. Pier P. Di Fiore, European Institute of Oncology, Milan, Italy (26). Fluorescence-activated cell sorting employing the 528 monoclonal anti-EGFR antibody was used to isolate a subline expressing high levels of this chimeric receptor. The 12CA5 anti-HA antibody was purchased from BAbCO, Berkeley, CA; the anti-InsR antibody (27) was a gift of Dr. Lynn Seely, University of California San Diego, La Jolla, CA, and the anti-phosphotyrosine antibody PY-20 was from Transduction Laboratories (Lexington, KY). A rabbit polyclonal anti-Ret antibody was raised against the 20-amino acid peptide (Lys-Arg-Arg-Asp-Tyr-Leu-Asp-Leu-Ala-Ala-Ser-Thr-Pro-Ser-Asp-Ser-Ile-Tyr-Asp) of the carboxyl terminus of the shorter splice isoform of Ret (28). The residue numbering in the present report follows that of Ret/ptc2. For comparison with residue numbering in c-Ret, see Durick *et al.* (29).

Synthesis and Purification of Peptides—Peptides were synthesized on an Advanced Chemtech MPS 350 at the Center for Molecular Genetics Peptide Synthesis Facility (University of California, San Diego) using Fmoc (*N*-(9-fluorenyl)methoxycarbonyl) chemistry. Peptides were dissolved in water and purified on a Sephadex G-15 column equilibrated in 0.12 M triethylammonium bicarbonate. Peak fractions (monitored at A_{274} or A_{280}) were dried in a speed vac (Savant), resuspended in water, and dried until the pH became neutral (total of four times). Peptides containing Trp were quantitated using A_{280} and an extinction coefficient of 5600 $M^{-1}cm^{-1}$. Tyr-containing peptides were quantitated using A_{274} and an extinction coefficient of 1400 $M^{-1}cm^{-1}$. The peptide lacking any absorptive amino acid was quantitated using amino acid analysis.

Mammalian Expression Vectors—The HA expression plasmid was constructed by subcloning a DNA fragment that encoded a Met initiation codon and an HA epitope tag into the mammalian expression vector pcDNA3 (Invitrogen) at the *Bam*HI and *Eco*RI sites of the polylinker region. cDNA clones corresponding to various regions of Enigma were excised from the pJG4-5 yeast expression vector using the restriction enzymes *Eco*RI and *Xho*I and fused in-frame to the HA epitope-tagged vector. The Ret/ptc2 and mutant Ret/ptc2 constructs in the RC/CMV expression plasmid (Invitrogen) were prepared as described elsewhere (29). All constructs were confirmed by sequencing (U. S. Biochemical Corp.)

Production of GST Fusion Proteins—Construction of plasmids expressing GST fusion proteins with the LIM domains of Enigma has been described (17). Two polymerase chain reaction primers were synthesized and used to isolate a DNA fragment encoding the carboxyl-terminal 61 amino acids (residues 536–596) of Ret. The polymerase chain reaction product bearing 5′ *Eco*RI and 3′ *Sal*I sites was cloned into the polylinker region pGEX-KG (Pharmacia Biotech Inc.). GST fusion proteins were prepared using standard procedures and the fusion proteins were immobilized on glutathione agarose beads (Sigma).

The Yeast Two-hybrid System—Reagents and procedures for the Lex-A based-yeast two-hybrid system were utilized (30). cDNAs coding wild type and mutant Ret/ptc2 were used to create LexA-Ret/ptc2 fusion proteins. cDNA fragments coding all the tested LIM domains were isolated by using *Pfu* polymerase and oligonucleotides that contained *Eco*RI and *Xho*I sites and fused in-frame to the pJG4-5 vector. Plasmids that direct the synthesis of LIM domains and full-length Enigma were introduced into EGY48/1840 yeast that contained different LexA fusion Ret/ptc2 constructs using the lithium acetate procedure. All constructs were confirmed by dideoxynucleotide sequencing, and the expression of fusion proteins of the appropriate size was confirmed by Western blotting with anti-HA antibody (BAbCO). β-Galactosidase activity was visualized on 5-bromo-4-chloro-3-indoyl β-D-galactoside galactose-containing plates or measured in solution according to Current Protocols of Molecular Biology (31). Yeast cells were grown in the presence of glucose or galactose and resuspended in an equal volume of buffer Z (100 mM $NaPO_4$, 10 mM KCl, 1 mM $MgSO_4$, 50 mM β-mercaptoethanol). Cells were diluted 1:10 or 1:20 with buffer Z and permeabilized by SDS and chloroform. *o*-Nitrophenyl β-*p*-galactopyranoside was used as a substrate and the reaction was stopped when a medium-yellow color had developed. Activity was calculated according to: units = 1000 (A_{420}/A_{600} *tv*) when *t* = time and *v* = volume (17).

Affinity Precipitation and Peptide Competition Assays—NIH3T3 cells overexpressing the EGFR/Ret chimera or EGFR were treated with 100 nM EGF for 10 min at 37 °C. Treated and untreated cells were lysed with a solution containing 50 mM Hepes (pH 7.4), 1% Triton X-100, 10% glycerol, 150 mM NaCl, 5 mM KCl, 1 mM EDTA, 2 mM phenylmethylsulfonyl fluoride, 1 mM Na_3VO_4, 10 mM benzamidine, 10 μg/ml of aprotinin and leupeptin. Rat1 cells (HIRC) overexpressing InsR (18) and 293 cells expressing Enigma or Ret/ptc2 were lysed in a similar fashion. The 293 cells were transfected with the HA-Enigma or deletion mutants using the calcium phosphate precipitation procedure (32). Expression of Ret/ptc2 and mutant Ret/ptc2 was accomplished similarly. Cells were harvested 48 h after transfection. The lysates were incubated with GST fusion proteins that were immobilized on glutathione agarose beads for 1 h at 4 °C with continuous agitation. For assay of peptide competition, the indicated concentrations of peptides were mixed with lysates prior to incubation with GST fusion proteins. Beads were then washed four times with lysis buffer. Material bound to beads was resolved by electrophoresis and probed with antibodies to Ret, InsR, Enigma, phosphotyrosine, or HA.

RESULTS

Identification of the Recognition Motif for LIM3 of Enigma—To determine the recognition motif for the LIM3 domain of Enigma, a random peptide library selection technique was used to study the consensus binding site. Random peptide library selection has been successfully used to determine the sequence specificity of the peptide-binding sites of SH2 and SH3 domains as well as the optimal substrates of protein kinases (33, 34). We constructed a fixed tyrosine peptide library comprising peptides of the sequence: Met-Ala-*X*-*X*-*X*-*X*-Tyr-*X*-*X*-*X*-*X*-Ala-Lys-Lys-Lys, where *X* indicates all amino acid except Trp, Cys, Ser, Thr, or Tyr. Trp and Cys were omitted to avoid problems with sequencing and oxidation. The total theoretical degeneracy of this library is 15^8. The Met-Ala sequence at the amino terminus provides two amino acids to verify that peptides from this mixture are being sequenced. Sequencing of these two residues also provides quantitation of the peptides present. Ala-12 provides a second quantitation and an estimate of how much peptide loss occurred during sequencing. The poly-Lys tail prevented wash-out during sequencing and improved the solubility of the mixtures.

LIM3 of Enigma was expressed as a GST fusion protein in *Escherichia coli*. The fusion protein was immobilized on glutathione agarose and incubated with the tyrosine peptide library. Unbound peptides were washed away and bound peptides were released by acid and subjected to micro-sequencing. The amino acids preferentially selected by LIM3 of Enigma at positions −4, −3, −2, −1 amino-terminal to the Tyr residue and +1, +2, +3, +4 carboxyl-terminal to the Tyr residue are shown in Fig. 1*A*. The greatest selectivity was observed at the −1 and +2 positions where Pro was preferred. At the −2 position, glycine was the preferred amino acid and at the +3 position, both Val and Ile were highly selected. Phe was preferred at position +4.

Because Pro was preferred at positions −1 and +2, a second library with the sequence Met-Ala-*X*-*X*-*X*-*X*-*X*-Pro-*X*-*X*-Pro-*X*-*X*-*X*-Ala-Lys-Lys-Lys in which Pro was fixed with two intervening amino acids was designed to further test selectivity. This library lacks only Cys and Trp and has a degeneracy of 18^{10}.

A

Tyrosine Random peptide Library

-4	-3	-2	-1	0	+1	+2	+3	+4
E(1.7)	E(1.7)	G(1.5)	P(2.0)	Y	G(1.3)	P(1.9)	V(2.4)	F(1.4)
I(1.4)	I(1.4)				H(1.3)		I(2.3)	H(1.4)
V(1.4)								

B

Proline Random peptide Library

-4	-3	-2	-1	0	+1	+2	+3	+4	+5
X	X	G(2.1)	P	M(1.6)	G(1.4)	P	I(1.9)	Y(1.7)	A(1.7)
				F(1.6)			M(1.5)	F(1.4)	
				Y(1.6)			V(1.3)		
				I(1.4)					

C

Consensus:	G	P	Hyd	G	P	Hyd	Y/F	A
Exon16/InsR:	G	P	L	G	P	L	Y	A

FIG. 1. **Substrate specificity of LIM3 of Enigma determined by binding degenerate peptide libraries.** Values in *parentheses* indicate the relative selectivity for the amino acids and *X* indicates no selectivity. The one-letter amino-acid code is used. LIM3 of Enigma was expressed as a GST fusion protein and immobilized on glutathione agarose. A Tyr-fixed random peptide library (*A*) or a Pro-fixed random peptide library (*B*) was presented to the immobilized GST-LIM3/Enigma. The unbound peptides were washed away and the retained peptide mixture was sequenced. The consensus peptide sequence is compared to the endocytic code of exon 16 of InsR (*C*).

This library also included Tyr at the 10 degenerate positions. The general motif determined by this library was similar to those found with the Tyr-fixed peptide library (Fig. 1*B*). In addition, a Tyr residue was highly selected at position +4. By comparison of these two motifs from two peptide library selections, the peptide sequence of Gly-Pro-Hyd-Gly-Pro-Hyd-Tyr/Phe-Ala was determined to be the recognition motif for LIM3 of Enigma (Fig. 1*C*). This peptide sequence is highly homologous to the sequence of exon 16 of InsR.

To confirm the binding motif for LIM3/Enigma, peptides were tested for their capacity to disrupt the complex of LIM3/Enigma with holo InsR. HIRC cell lysates containing InsR were incubated with GST-LIM3/Enigma without or with competitor peptides. As shown in Fig. 2, binding of InsR to GST-LIM3/Enigma was inhibited by a 12-amino acid peptide (Asp-Gly-Pro-Leu-Gly-Pro-Leu-Tyr-Ala-Ser-Ser-Asn) corresponding to exon 16 of the InsR but not by mutant peptides (Asp-Gly-Pro-Leu-**Ala**-Pro-Leu-**Ala**-Ala-Ser-Ser-Asn and Asp-Gly-**Ala**-Leu-Gly-**Ala**-Leu-Tyr-Ala-Ser-Ser-Asn). The single substitution of the Leu immediately preceding the Tyr for Ile did not affect its ability to compete for InsR binding, confirming the random peptide library selection of a hydrophobic residue at position +3. These peptide competition results demonstrate that both the Pro and Tyr residues are required to mediate interaction of LIM3 of Enigma with exon 16 of InsR.

Differential Recognition of Ret/ptc2 and InsR by LIM Domains of Enigma—When an oncogenic form of Ret, Ret/ptc2, was used in a yeast two hybrid screen to identify interacting proteins, several SH2 domain containing proteins and Enigma were isolated (25). To determine the domains of Enigma responsible for this interaction with Ret/ptc2, regions of Enigma were cloned into pJG4-5 and tested for their ability to bind Ret/ptc2 expressed as a LexA fusion protein in pEG202 in yeast. As shown in Fig. 3*A*, full-length Enigma bound similarly to Ret/ptc2 and exon 16/InsR. The amino-terminal 279 amino acids of Enigma did not interact with either Ret/ptc2 or exon 16/InsR. The carboxyl-terminal 275 amino acids containing the three LIM domains were thus responsible for protein-protein interactions with both Ret and InsR. When the three LIM domains of Enigma were divided into individual LIM domains, LIM2 of Enigma bound Ret/ptc2 but not exon 16/InsR. Conversely, LIM3 of Enigma bound exon 16/InsR but not Ret/ptc2. These results demonstrate that LIM2 of Enigma was responsible for the association of Enigma with Ret/ptc2 and could be physically separated from LIM3 of Enigma which was responsible for Enigma association with InsR. LIM1 of Enigma, which bound two atoms of zinc characteristic of LIM domains[3] did not associate with either RET/ptc2 or InsR.

The specificity of interaction of Ret/ptc2 and InsR with other LIM domains was further examined. As shown in Fig. 3*B*, LIM domains of Mec-3, Isl-1, Lmx-1, zyxin, CRP, and paxillin did not recognize Ret/ptc2 or exon 16/InsR. The specificity for recognition thus resides in the LIM domains of Enigma.

Mapping the Interaction Site of of Ret with Enigma—Deletion of the carboxyl terminus of Ret/ptc2 distal to the conserved tyrosine kinase core abolished both mitogenic activity and Enigma binding (25). To determine whether the carboxyl-terminal region of Ret is sufficient to support the interaction, the carboxyl-terminal 61 amino acids of Ret/ptc2 (residues 536 to 596) were expressed as a GST fusion protein (GST-C′/Ret) and tested for their ability to bind Enigma. GST and GST-C′/Ret were immobilized on glutathione agarose and mixed with full-length, amino-terminal or carboxyl-terminal domains of Enigma that were expressed as HA epitope-tagged fusion proteins in 293 cells (Fig. 4*A*). Equal amounts of GST and GST-C′/Ret were assessed for their ability to bind these Enigma proteins (Fig. 4*B*). GST-C′/Ret bound full-length Enigma and

[3] D. Winge, personal communication.

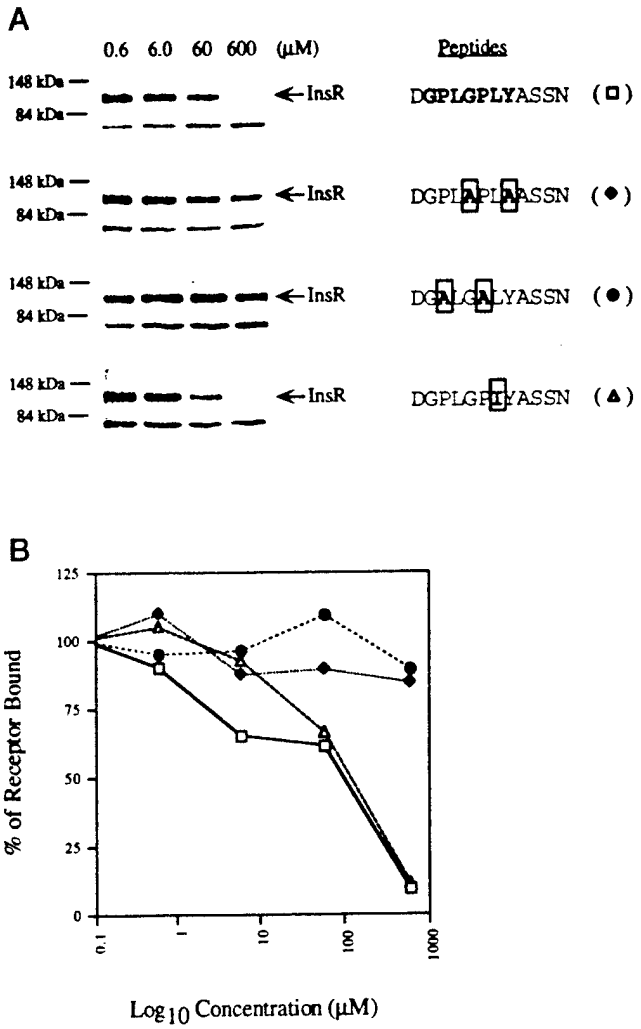


FIG. 2. **Peptide competition with InsR binding to LIM3 of Enigma (*A*) lysates of Rat1 cells that overexpressed InsR were incubated with GST-LIM3/Enigma in the presence of varying concentrations of the indicated peptides.** GST beads were washed and the associated InsR eluted and detected by Western blotting using an antibody specific to the β subunit of InsR. *B*, bound receptor was quantitated by scanning densitometry and plotted against the concentration of competitor peptide. The amount bound in the absence of competitor peptide was set at 100%.

the carboxyl terminus containing the three LIM domains of Enigma but failed to bind the amino terminus of Enigma. There was no binding of any of the three forms of Enigma protein to GST. These results indicate that the carboxyl-terminal 61 amino acids of Ret contain the motif that is required for the association of Ret with Enigma.

To map the binding site within the carboxyl terminus of Ret, peptides were used to specifically disrupt the complexes of Ret with Enigma. Four 20-amino acid peptides (Fig. 5*A*) were assayed for their ability to disrupt the interaction of Ret with Enigma. Full-length Enigma was mixed with the indicated peptides and interacted with immobilized GST-C′/Ret. As shown in Fig. 5B, incubation with wild type peptide, Asn-Lys-Leu-Tyr, and mutant peptide, Asn-Lys-Leu-**Phe**, were sufficient to displace Enigma binding to the carboxyl terminus of Ret, while the Ret mutant peptide **Ala**-Lys-Leu-**Ala** and the peptide corresponding to exon 16/InsR were without effect. These results indicate that the sequence **Asn-Lys-Leu-Tyr** in the carboxyl terminus of Ret is necessary for interaction with the LIM2 domain of Enigma. Mutation of the Asn and Tyr residues abolished interaction but a Phe substitution for Tyr was tolerated.

Specificity was further assessed by using these peptides to disrupt interaction of an EGFR/Ret chimera (26) with immobilized GST-LIM2 of Enigma. NIH3T3 cell lysates expressing EGFR/Ret were incubated with GST-LIM2/Enigma in the presence of the indicated peptides, and bound receptors were detected by an anti-Ret antibody. Similar to the results in Fig. 5*B*, peptides Asn-Lys-Leu-Tyr and Asn-Lys-Leu-**Phe** displaced EGFR/Ret binding to LIM2/Enigma (Fig. 5*C*). However, the peptide with the two amino acid mutation to **Ala**-Lys-Leu-**Ala** and the exon 16/InsR peptide failed to compete for the binding, confirming the Asn-Lys-Leu-Tyr sequence at the carboxyl terminus of Ret is the core recognition site for LIM2/Enigma. EGFR alone did not interact with Enigma (data not shown).

Although the mutant peptide Asn-Lys-Leu-**Phe** blocked the interaction of Enigma and Ret using GST-fusion protein assays, mutation of Tyr^{586} to Phe in Ret/ptc2 decreased this interaction as assayed in a yeast two hybrid system and decreased Ret/ptc2-stimulated DNA synthesis in microinjection experiments in mouse fibroblasts (25). To clarify these differing results, the effects of replacement of Tyr^{586} with Phe in Ret/ptc2 on Enigma binding were quantitated. Wild-type Ret/ptc2 (Tyr^{586}) and mutant Y586F Ret/ptc2 (Phe^{586}) were expressed in 293 cells and the relative affinities of these proteins for the LIM domains of Enigma were measured. As shown in Fig. 5*D* the affinity of Tyr^{586} exceeded that of Y586F Ret/ptc2 for Enigma by approximately 5-fold. Deletion of the carboxyl terminus containing this region, *i.e.* Ret/ptc2 truncated at residue 574 completely abolished the interaction (data not shown). LIM2 of Enigma thus recognized the Phe substituted carboxyl terminus of Ret but with lower affinity compared to wild type Ret/ptc2 with the Tyr-containing sequence. The differing results using the yeast two hybrid system and *in vitro* peptide competition are explained by the decreased affinity of Y586F compared to wild type Ret/ptc2 for LIM 2 of Enigma.

Tyrosine Phosphorylation of Ret Is Not Required for Binding to LIM2 of Enigma—To investigate whether tyrosine kinase activation was required for the association between LIM2 Enigma and Ret/ptc2, the EGFR/Ret chimeric protein was used. Because the ligand for the Ret tyrosine kinase receptor is unknown, the chimera generated by fusing the extracellular and transmembrane domains of EGFR and the intracellular domain of Ret was used (Fig. 6*A*). EGF activated the Ret tyrosine kinase activity and mitogenic responses of this chimera (26). NIH3T3 cells overexpressing EGFR/Ret were treated without or with EGF and cell lysates were mixed with GST or GST-LIM domains of Enigma. As shown in Fig. 6*B*, only GST-LIM2/Enigma interacted with EGFR/Ret. GST, GST-LIM1, or GST-LIM3 of Enigma did not bind. Ligand treatment did not effect the binding of EGFR/Ret to LIM2 of Enigma (*left panel*). Phosphorylation of the EGFR/Ret chimera was also examined by Western blotting using an anti-phosphotyrosine antibody. GST-LIM2 of Enigma interacted with phosphorylated as well as unphosphorylated EGFR/Ret receptors. The interaction of LIM2 of Enigma with Ret is thus independent of ligand activation and Ret autophosphorylation.

DISCUSSION

The growing number of proteins that contain one or more LIM domains function in a variety of pathways and locations within the cell, implicating LIM domains as versatile protein modules that are capable of acting in diverse cellular contexts. Although the NMR structure of LIM2 of CRP resembles the DNA binding domain of the GATA-1 transcription factor, no direct evidence that a LIM domain binds to nucleic acids has been presented. Indeed, a lack of affinity for target DNA sequences has been reported for the LIM domains of Mec-3 (35).

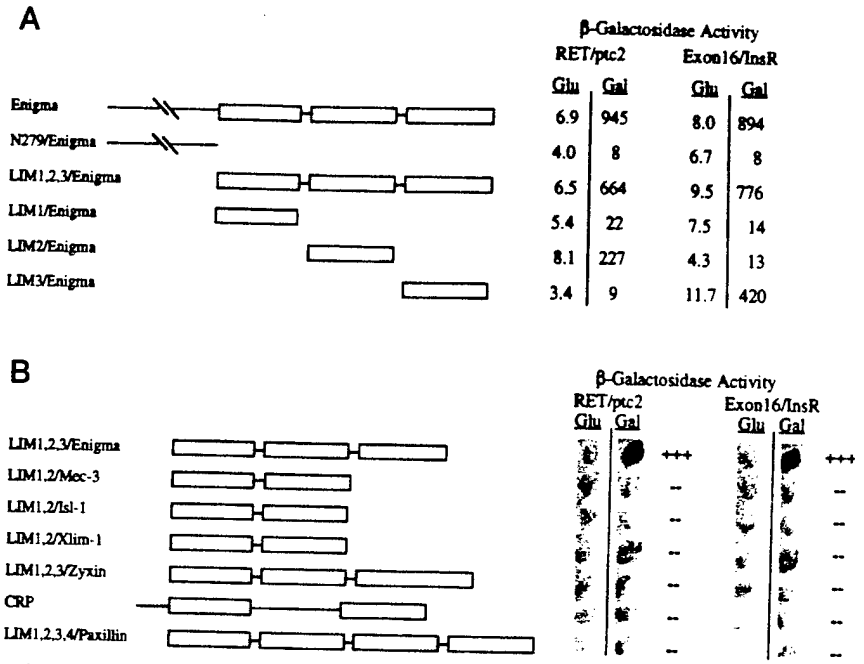


FIG. 3. **Specific interaction of LIM domains of Enigma with Ret/ptc2 and exon 16/InsR in a yeast two hybrid system.** *A*, differential recognition of Ret/ptc2 and exon 16/InsR by LIM2 and LIM3 of Enigma. *B*, comparison of Ret/ptc2 and exon 16/InsR interaction with LIM domains of Enigma and other proteins. Ret/ptc2 and exon 16 InsR were expressed as LexA DNA binding domain fusion proteins in plasmid pEG202. β-Galactosidase assays were performed on yeast expressing individual LexA fusion proteins and the indicated B42 activation domain fusion proteins in plasmid pJG4-5. β-Galactosidase activity of each transformant was visualized and measured in solution. "+" indicates dark blue and "–" indicates white colonies. Numbers indicate β-galactosidase activity units quantitated from solution assays.

Most available evidence indicates that LIM domains function in specific protein-protein interactions (1, 8).

The present studies demonstrate that LIM2 of Enigma specifically interacts with Ret while LIM3 of Enigma specifically interacts with InsR. The Asn-Lys-Leu-Tyr sequence at the carboxyl terminus of Ret was essential for the formation of the Ret-Engima complex. For the interaction of InsR with Enigma, the Gly-Pro-Leu-Gly-Pro-Leu-Tyr sequence of the juxtamembrane region of InsR was required. Both LIM2 and LIM3 recognized Tyr-containing motifs located outside of the tyrosine-kinase cores of Ret and InsR. Although the recognition motifs for LIM2 and LIM3 of Enigma share sequence similarity, they were not exchangeable, demonstrating that the two LIM domains have the ability to distinguish between two Tyr-based motifs. These results also indicate that individual LIM domains within a single protein have distinct partner preferences. Because the structural features of LIM domains are highly conserved, sequences other than the conserved residues that are involved in metal coordination must be important for defining the selectivity of individual LIM domains for their particular partner. LIM1 of Enigma failed to recognize either Ret or InsR and is likely to have a yet unidentified target protein in cells.

Tyr-based motifs serve a number of functions. Tyr-containing tight turns are the essential structural feature of the endocytic codes of many proteins (19). Four to six amino acid sequences containing an essential Tyr residue also function as lysosomal and trans-Golgi to basolateral surface targeting codes (36–38). There is no evidence for covalent modification of Tyr residues within these trafficking codes. Phosphorylated Tyr residues in specific sequence contexts serve as the recognition motif for SH2 and PTB domains (33, 39, 40). The present studies indicate that Tyr-containing sequences also function as recognition elements for certain LIM domains. The Gly-Pro-Leu-Tyr motif of InsR that is recognized by LIM3 of Enigma forms a Tyr-containing tight turn (20). Use of two random

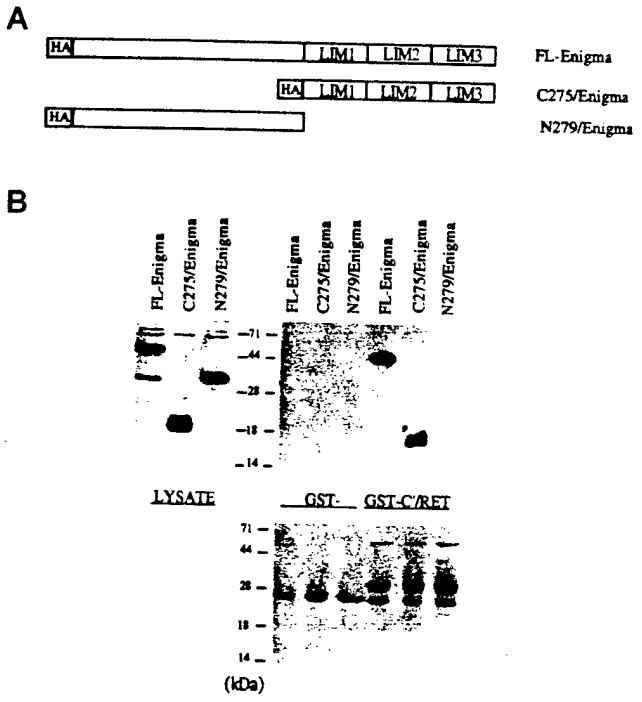
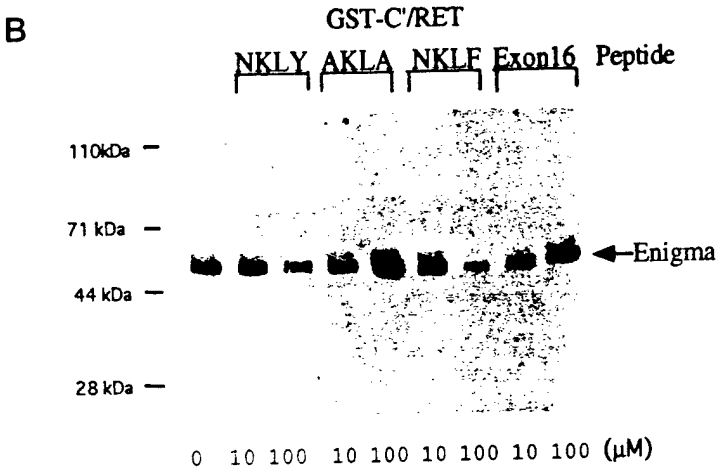
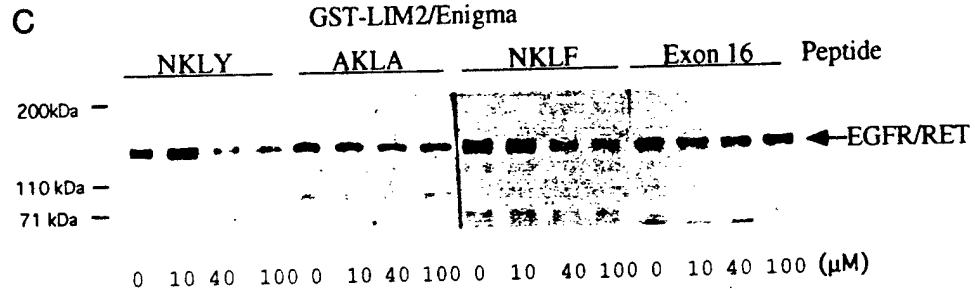


FIG. 4. **Direct association of Enigma with the carboxyl terminus of Ret.** *A*, schematic representation of the Enigma fragments used in the binding assays. Full-length and regions of Enigma were expressed as HA epitope-tagged fusion proteins. *B*, binding of Enigma to the carboxyl-terminal 61 amino acids (residues 536–596) of Ret. 293 cells were transfected with the indicated expression vectors; cell lysates were prepared 48 h later and incubated with GST or GST-C'/Ret. Bound material was analyzed by Western blotting with antibody to the HA epitope (*upper*). The *bottom panel* is a Coomassie Blue-stained gel to quantitate GST and GST fusion protein used in the assays. The *left panel* shows the amount of proteins in the lysate prior to interaction with GST C'/Ret.

A

Peptide	Amino Acid Sequence
NKLY	PSTWIENKLYGRISHAFTRF
AKLA	PSTWIEAKLAGRISHAFTRF
NKLF	PSTWIENKLFGRISHAFTRF
Exon16	DGPLGPLYASSNPEYLSASD





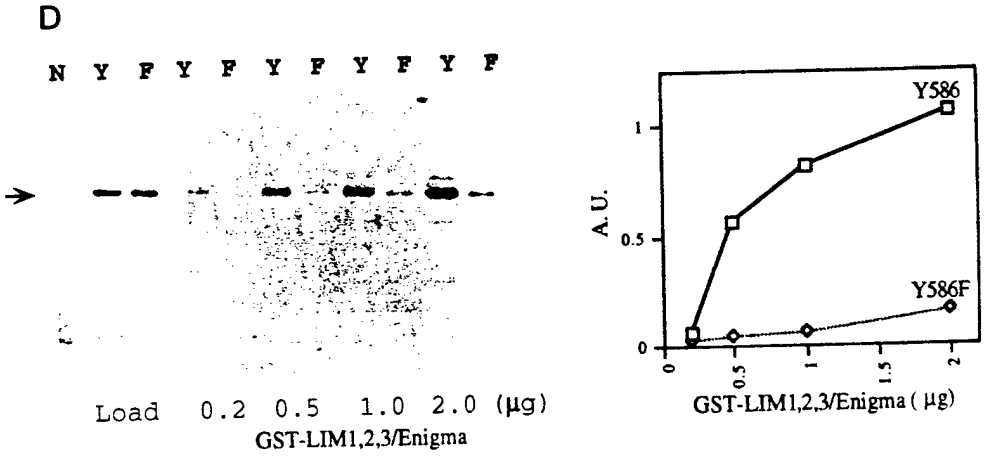


FIG. 5. **Sequence specificity of Ret binding to Enigma using peptide competition assays and GST-fusions.** *A*, amino acid sequences of the peptides used for competition. The NKLY peptide corresponds to the carboxyl-terminal 20 amino acids (residues 577–596) of Ret/ptc2. Exon 16 corresponds to the juxtamembrane region of InsR. *B*, specific dissociation of Enigma and GST-C'/Ret complexes by peptides. The HA-tagged Enigma protein was expressed in 293 cells and mixed with GST-C'/Ret without or with the indicated concentrations of peptides. Bound Enigma protein was detected by anti-HA antibody. *C*, peptide competition of EGFR/Ret binding to GST-LIM2 of Enigma. Lysates were prepared from NIH3T3 cells expressing the EGFR/Ret chimera and bound receptor was detected using the anti-Ret antibody. *D*, comparison of binding of Tyr^{586} with Y586F Ret/ptc2 to the carboxyl terminus of Enigma. Wild-type Ret/ptc2 (Tyr^{586}) and the point mutation Y586F Ret/ptc2 (Phe^{586}) were expressed in 293 cells. Equal amounts were mixed with the indicated amount of GST-LIM1,2,3/Enigma, and bound Ret/ptc2 was detected by the anti-Ret antibody (*left*). Bound protein was quantitated by scanning densitometry and plotted against the amount of GST LIM domains of Enigma used (*right*). The amount bound is expressed as absorbance units (*A.U.*)

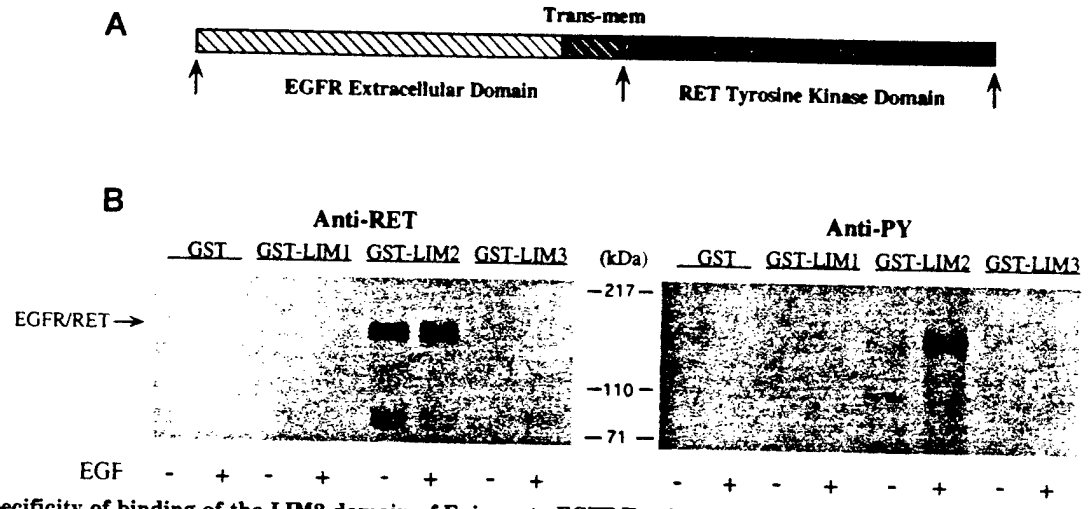


FIG. 6. **Specificity of binding of the LIM2 domain of Enigma to EGFR/Ret *in vitro*.** *A*, schematic structure of the EGFR/Ret chimera indicating the extracellular and transmembrane domains of EGFR and the intracellular domain of Ret. The chimeric receptor utilized EGF as a ligand to activate its tyrosine kinase. *B*, immunoblot analysis of the binding of EGFR/Ret chimera by GST fusion proteins. NIH3T3 cells overexpressing the EGFR/Ret chimera were treated without or with 100 nM EGF. Cell lysates were incubated with GST or GST-LIM domains of Enigma. Bound proteins were analyzed by Western blotting with antibodies to Ret or to phosphotyrosine.

peptide libraries indicated that the target recognized by LIM3 of Enigma consisted of the more extended sequence Gly-Pro-Leu-Gly-Pro-Leu-Tyr-Ala. The Asn-Lys-Leu-Tyr motif of Ret that is recognized by LIM2 of Enigma resembles the endocytic sequence in the LDL receptor which forms a Tyr-containing tight turn (41).

The function of the target sequences in Ret and InsR are different. LIM3 of Enigma recognized the major endocytic code of InsR while LIM2 of Enigma interacted with the carboxyl terminus of Ret. Ret is a protein tyrosine kinase receptor implicated in several disease processes. Mutations that inactivate its tyrosine kinase result in Hirschsprung's disease characterized by defective sympathetic innervation of the large intestine (42, 43). Activating mutations characterize a group of inherited multiple endocrine neoplasia type syndromes that include MEN2A, MEN2B, and familial medullary thyroid cancer (44, 45). Gene rearrangements including the one fusing the type 1 regulatory subunit of cyclic AMP-dependent protein kinase to the tyrosine kinase domain of Ret (Ret/ptc2) occur as oncogenic events in papillary thyroid carcinoma (28). In a nuclear microinjection assay the mitogenic activity of Ret/ptc2 was abolished by carboxyl-terminal truncation to residue 574 or by the mutation Y586F (25, 29). Deletion of sequences distal to residue 574 abolished interaction with Enigma and mutation of Tyr^{586} reduced the affinity of Enigma for Ret. Additionally, co-expression of the LIM domains of Enigma blocked the mitogenic activity of Ret/ptc2, implicating involvement of Enigma in the mitogenic signaling of Ret (25). The ability of LIM3 of Enigma to recognize the active endocytic codes of InsR fulfills the first property of the endocytic mechanism, but additional functional criteria will be necessary to critically test the hypothesis that Enigma functions in endocytosis of InsR.

Interactions of LIM domains of Enigma with these two receptors did not require either tyrosine kinase activity or tyrosine autophosphorylation on their target sequences. However, for both endocytosis of InsR and mitogenic signaling by Ret, activation of tyrosine kinase activity is necessary. Interactions of Enigma with InsR and Ret are thus proposed to be necessary but not sufficient to support these biological processes. Interestingly, Tyr^{1062} in holo Ret corresponding to Tyr^{586} in Ret/ptc2 is phosphorylated when expressed in COS cells (46). The stoichiometry was not determined but this could provide a mechanism for the reversible association of Ret with Enigma.

Most LIM proteins contain more than one LIM domain (1, 8). Finding distinct targets for two LIM domains of a single protein not only supports the hypothesis that LIM domains function in protein-protein interactions but indicates a possible adaptor function to assemble multiple proteins into a complex. Alternatively, multiple LIM domains could function to differentially assemble proteins with distinct receptors. The finding that binding of InsR and of Ret to Enigma does not require the protein tyrosine kinase activity of either receptor distinguishes LIM domain interactions from those of SH2 and PTB domains which direct assembly that is dependent on tyrosine kinase activity and covalent modifications of proteins (33, 39). Because the processes requiring the target sequences in InsR and Ret that are recognized by Engima do depend on the tyrosine kinase activity of these receptors, mechanisms of function of the assembled LIM domain complexes must coordinate with tyrosine kinase activity perhaps via substrate phosphorylation or protein assemblies.

Acknowledgments—The authors thank P. Paolo Di Fiore for cells expressing the EGFR/Ret chimera and M. Pierotti for the Ret/ptc2 clone.

REFERENCES

1. Gill, G. N. (1995) *Structure* **3**, 1285–1289
2. Michelsen, J. W., Schmeichel, K. L., Beckerle, M. C., and Winge, D. R. (1993) *Proc. Natl. Acad. Sci. U. S. A.* **90**, 4404–4408
3. Perez-Alvarado, G. C., Miles, C., Michelsen, J. W., Louis, H. A., Winge, D. R., Beckerle, M. C., and Summers, M. F. (1994) *Nat. Struct. Biol.* **1**, 388–398
4. Freyd, G., Kim, S. K., and Horvitz, H. R. (1990) *Nature* **344**, 876–879
5. Karlsson, O., Thor, S., Norberg, T., Ohlsson, H., and Edlund, T. (1990) *Nature* **344**, 879–882
6. Way, J. C., and Chalfie, M. (1988) *Cell* **54**, 5–16
7. Taira, M. M., Jamrich, M., Good, P. J., and Dawid, I. B. (1992) *Genes Dev.* **6**, 356–366
8. Dawid, I. B., Toyama, R., and Taira, M. (1995) *C. R. Acad. Sci. de la Vie/Life Sci. (Paris)* **318**, 295–306
9. Sadler, I., Crawford, A. W., Michelsen, J. W., and Beckerle, M. C. (1992) *J. Cell Biol.* **119**, 1573–1587
10. Turner, C. E. (1994) *BioEssays* **16**, 47–52
11. Crawford, A. W., Pino, J. D., and Beckerle, M. C. (1994) *J. Cell Biol.* **124**, 117–127
12. Boehm, T., Greenberg, J. M., Buluwela, L., Lavenir, I., Forster, A., and Rabbitts, T. H. (1990) *EMBO J.* **9**, 857–868
13. Royer-Pokora, B., Loos, U. Y., and Ludwig, W. D. (1991) *Oncogene* **6**, 1887–1893
14. Mizuno, K., Okano, I., Ohashi, K., Nunoue, K., Kuma, K., Miyata, T., and Nakamura, T. (1994) *Oncogene* **9**, 1605–1612
15. Schmeichel, K. L., and Beckerle, M. C. (1994) *Cell* **79**, 211–219
16. Feuerstein, R., Wang, X., Song, D., Cooke, N. E., and Liebhaber, S. A. (1994) *Proc. Natl. Acad. Sci. U. S. A.* **91**, 10655–10659
17. Wu, R. Y., and Gill, G. N. (1994) *J. Biol. Chem.* **269**, 25085–25090
18. Rajagopalan, M., Neidigh, J. L., and McClain, D. A. (1991) *J. Biol. Chem.* **266**, 23068–23073
19. Trowbridge, I. S., Collawn, J. F., and Hopkins, C. R. (1993) *Annu. Rev. Cell*

Biol. **9,** 129–161
20. Backer, J. M., Shoelson, S. E., Weiss, M. A., Hua, Q. X., Cheatham, R. B., Harig, E., Cahill, D. C., and White, M. F. (1992) *J. Cell Biol.* **118,** 831–839
21. Chang, C.-P., Lazar, C. S., Walsh, B. J., Komuro, M., Collawn, J. F., Kuhn, L. A., Tainer, J. A., Trowbridge, I. S., Farquhar, M. G., Rosenfeld, M. G., Wiley, H. S., and Gill, G. N. (1993) *J. Biol. Chem.* **268,** 19312–19320
22. German, M. S., Wang, J., Chadwick, R. B., and Rutter, W. J. (1992) *Genes Dev.* **6,** 2165–2176
23. Wadman, I., Li, J., Bash, R. O., Forster, A., Osada, H., Rabbitts, T. H., and Baer, R. (1994) *EMBO J.* **13,** 4831–4839
24. Bach, I., Rhodes, S. J., Pearse, R. V., II, Heinzel, T., Gloss, B., Scully, K. M., Sawchenko, P. E., and Rosenfeld, M. G. (1995) *Proc. Natl. Acad. Sci. U. S. A.* **92,** 2720–2724
25. Durick, K., Wu, R.-Y., Gill, G. N., and Taylor, S. S. (1996) *J. Biol. Chem.* **271,** 12691–12694
26. Santoro, M., Wong, W. T., Aroca, P., Santos, E., Matoskova, B., Grieco, M., Fusco, A., and di Fiore, P. P. (1994) *Mol. Cell. Biol.* **14,** 663–675
27. Soos, M. A., Field, C. E., Lammers, R., Ullrich, A., Zhang, B., Roth, R. A., Andersen, A. S., Kjeldsen, T., and Siddle, K. (1992) *J. Biol. Chem.* **267,** 12955–12963
28. Bongarzone, I., Monzini, N., Borrello, M. G., Carcano, C., Ferraresi, G., Arighi, E., Mondellini, P., Della Porta, G., and Pierotti, M. A. (1993) *Mol. Cell. Biol.* **13,** 358–366
29. Durick, K., Yao, V. J., Borrello, M. G., Bongarzone, I., Pierotti, M. A., and Taylor, S. S. (1995) *J. Biol. Chem.* **270,** 24642–24645
30. Gyuris, J., Golemis, E., Chertkov, H., and Brent, R. (1993) *Cell* **75,** 791–803
31. Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., and Struhl, K. (1994) *Current Protocols of Molecular Biology*, p. 113.14.1, John Wiley & Sons, New York
32. Chen, C., and Okayama, H. (1987) *Mol. Cell. Biol.* **7,** 2745–2752
33. Songyang, Z., Shoelson, S. E., Chaudhuri, M., Gish, G., Pawson, T., Haser, W. G., King, F., Roberts, T., Ratnofsky, S., Lechleider, R. J., Neel, B. G., Birge, R. B., Fajardo, J. E., Chou, M., Hanafusa, H., Schaffhausen, B., and Cantley, L. C. (1993) *Cell* **72,** 767–778
34. Songyang, Z., Blechner, S., Hoagland, N., Hoekstra, M. F., Piwnica-Worms, H., and Cantley, L. C. (1994) *Curr. Biol.* **4,** 973–982
35. Xue, D., Tu, Y., and Chalfie, M. (1993) *Science* **261,** 1324–1328
36. Johnson, K. F., and Kornfeld, S. (1992) *J. Cell Biol.* **119,** 249–257
37. Letourneur, F., and Klausner, R. D. (1992) *Cell* **69,** 1143–1157
38. Matter, K., Yamamoto, E. M., and Mellman, I. (1994) *J. Cell Biol.* **126,** 991–1004
39. Pawson, T. (1995) *Nature* **373,** 573–580
40. Kavanaugh, W. M., Turck, C. W., and Williams, L. T. (1995) *Science* **268,** 1178–1179
41. Bansal, A., and Gierasch, L. M. (1991) *Cell* **67,** 1195–1201
42. Romeo, G., Ronchetto, P., Luo, Y. L., Barone, V., Seri, M., Ceccherini, I., Pasini, B., Bocciardi, R., Lerone, M., Kaariainen, H., and Martucciello, G. (1994) *Nature* **367,** 377–378
43. Edery, P., Lyonnet, S., Mulligan, L. M., Pelet, A., Dow, E., Abel, L. Holder, S., Nihoul-Fékété, C., Poner, B. A. J., and Munnich, A. (1994) *Nature* **367,** 378–383
44. Mulligan, L. M., Kwok, J. B. J., Healey, C. S., Elsdon, M. J., Eng, C., Gardner, E., Love, D. R., Mole, S. E., Moore, J. K., Papi, L., Ponder, M. A., Telenius, H., Tunnacliffe, A., and Ponder, B. A. J. (1993) *Nature* **363,** 458–376
45. Hofstra, R. M. W., Landsvater, R. M., Ceaccherini, I., Stulp, R. P., Stelwagen, T., Luo, Y., Pasini, B., Höppener, J. W. M., van Amstel, H. K. P., Romeo, G., Lips, C. J. M., and Buys, C. H. C. M. (1994) *Nature* **367,** 375–376
46. Liu, Y., Vega, Q. C., Decker, R. A., Pandey, A., Worby, C. A., and Dixon, J. E. (1996) *J. Biol. Chem.* **271,** 5309–5312